

Will crypto-nuclear powers come clean?

Next week's crucial meeting of the parties to the Non-Proliferation Treaty cannot put off some consideration of what will happen to countries believed to have nuclear weapons up their sleeves.

WHEN is a country a nuclear power? "When it has nuclear weapons, of course!" So what is to be made of Israel, which has operated a reactor at Davona in the Negev Desert since the early 1960s, which has declined to sign the Nuclear Non-Proliferation Treaty (NPT), which has jailed for life a technician who told a British newspaper of an underground plutonium factory in the desert, but which seems not to have carried out a nuclear test? Or of India, which conducted an underground nuclear explosion in 1974, but which insisted that its interests were exclusively in peaceful uses of nuclear explosives, which has also declined to sign the NPT, but which has not carried out a nuclear test for 20 years?

The conventional description is that they are 'cryptic' nuclear powers; they have not advertised their possession of nuclear weapons, but their neighbours must in prudence behave as if they had at least modest numbers. (Pakistan is known to have responded to India's status with at least a laboratory-scale programme of development in plutonium-separation. Iraq seems to have responded on behalf of Israel's Muslim neighbours with a nuclear development programme, now dismantled; Iran may have taken over where Iraq had to leave off.) It is to the credit of the President of Egypt, Hosni Mubarak, that he has raised the issue of the crypto-nuclear powers on the eve of the fourth and crucial review conference of the NPT in Geneva next week.

This meeting is crucial because the NPT will lapse this year unless the parties to the treaty collectively agree that it should continue. Against the odds, and despite the tardiness with which the nuclear powers recognized that this long-standing date on their calendars would come round, the prospects for the conference are more cheerful than they might have been. It seems assured that there will be enough votes to keep the NPT alive; the outstanding issue is whether, like the version ratified in 1970, it will be a fixed-

term treaty. This has been primarily accomplished by the willingness of the principal nuclear powers to negotiate among themselves a treaty to ban the testing of nuclear weapons for the rest of time. The promised negotiations may yet hit unexpected snags, but the promise is a dramatic capstone on the network of bilateral arms control treaties between Russia and the United States negotiated in the past few years. If it had been made five years ago, a great deal of trouble would have been saved.

And now, of course, there are the crypto-nuclear powers to be brought within the fold. Egypt's dialogue with Israel is central to this task. Israel has a strong case for being a nuclear power. For much of its history, its neighbours have denied its right to exist. What better means of deterring a concerted attack than implicitly to threaten enemies' cities with destruction? Even what is now called the Middle East peace process has not ended that sad state of affairs. The delicacy of Egypt's approach to Israel shows an understanding of this point, but there is no obvious solution in sight. Formal external military guarantees would not suffice. But on present form, it will be a long time before the peace process has gone far enough to be convincing. And while a solution of the problem would be to admit Israel to the NPT with the status of a nuclear power, that would simply invite North Korea to apply for membership in that guise.

In the end, the only solution is a general understanding in the United Nations that the manufacture of nuclear weapons is unacceptable, inviting all necessary sanctions to bring it to a halt. But the established nuclear powers (China included) will have to go much further than the promise of a test-ban before that is feasible. A verifiable cut-off of the production of fissile material for military purposes is the next step. It would be best that the NPT should not be made a perpetual treaty at least until that has been achieved. □

nature

**New
address**

From 18 April 1995 our address will be:

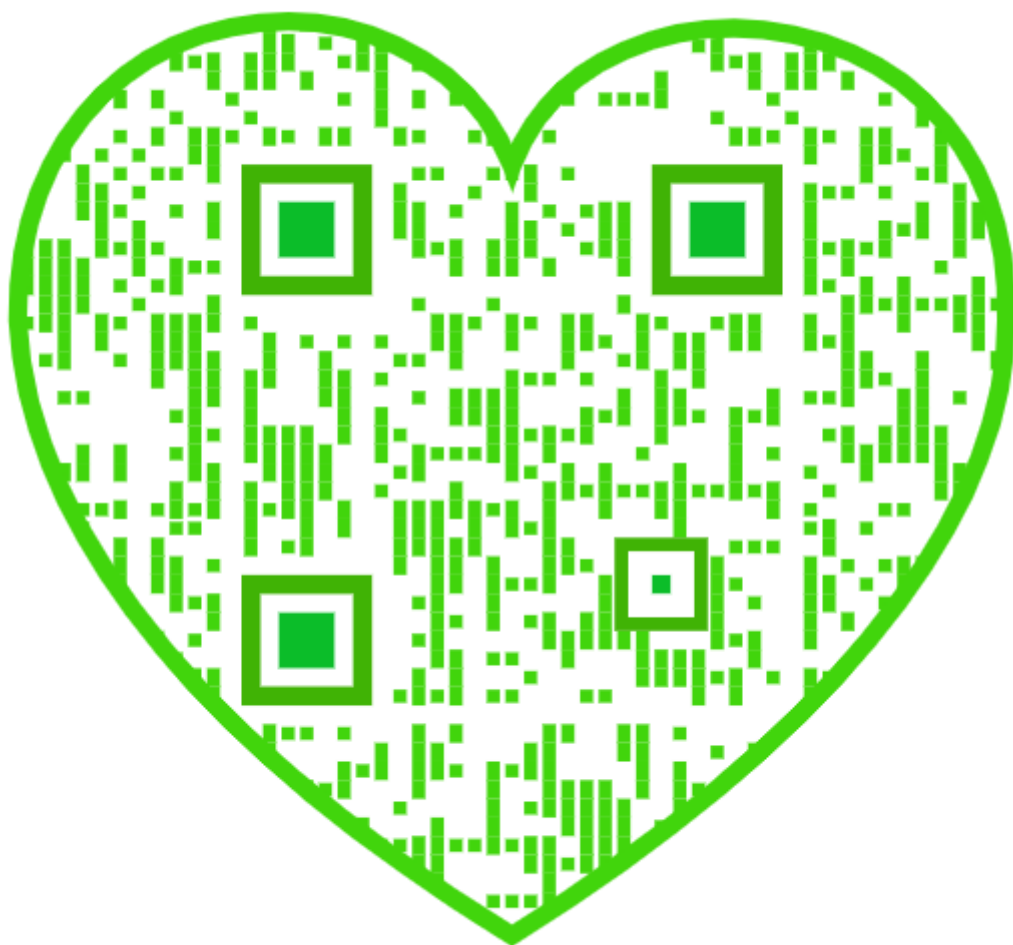
Porters South
Crinan Street
London N1 9SQ

Tel: +44 (0)171 833 4000
Fax: +44 (0)171 843 4596/7

Give CSIRO a rest!

Australia's research organization needs a period without formal inquiry into its ways of working.

AUSTRALIA'S delight in sports of all kinds sometimes gets the better of it. The sport of inquiring into the condition of the country's chief research organization, the Commonwealth Industrial and Scientific Research Organization (CSIRO), of making recommendations and causing a little administra-



tive chaos as a consequence, is especially at risk of getting out of hand. External inquiries seem to follow each other at intervals of five years or so. The latest development is the publication (last week) of an internal report on the management of CSIRO by a subcommittee of its management board that recommends, in effect, the unscrambling of an earlier reorganization (see page 587). Quite when the new arrangements will be in place is not, at this stage, clear; the board is still searching for a chief executive to run the organization as a whole.

There are two reasons why CSIRO is repeatedly the target of inquiries. One is that Australia is over-governed. There are more politicians in charge of some portfolio of government business, at the level of the states as well as the federal government, than there are urgent and soluble problems for them to tackle. Inquiring into the condition of CSIRO is a natural occupation for the underemployed, but sheer size also makes CSIRO a natural object of curiosity. So, too, does the culture of the organization, which still reflects something of the complacency of the days when CSIRO was virtually synonymous with research in Australia. Now, several universities have reputations in the field that command equal respect.

Nevertheless, there is a case for giving CSIRO a respite from public inquiry. In the past few decades, it has navigated the transition from the comprehensively avuncular research organization it once was, believing that it knew better than others what research Australia needed, to one that recognizes that it must deliver useful intellectual goods to customers who are most often industrial companies (or groupings thereof) or government departments. One of the distinctive features of the organization is that many of its people's first interests lie in basic research. To the extent that they have adapted to the idea that their customers' interests are paramount, they collectively have the opportunity to provide Australia with a practically orientated research organization more alert than would otherwise be the case to the importance of basic research. They should be given a chance to work that out for themselves.

Whether the details of the proposed management reorganization will prove sufficient is another matter. The chief change proposed is that the grouping of 'divisions' (free-standing laboratories) into 'institutes' with similar interests should be abolished. Among other things, that should rid the organization of some of the bureaucracy about which there have been endless complaints. (Some academic researchers say they cannot afford to use CSIRO services because they are too expensive; individual laboratories may be able to be more accommodating.) But then it is proposed that the programmes of individual laboratories should evolve from regular meetings of like-minded laboratory heads and people from headquarters. Much will depend on the spirit in which these discussions take place. How flexible will the bosses be (or be allowed by budgetary constraints to be)? The importance of public accountability notwithstanding, there is a strong case that CSIRO should have the freedom to work out the future pattern pragmatically, case by case. □

Russian genome project

Responsibility for Russia's modest interest in the human genome may be characteristically politicized.

RUSSIA has plenty on its mind just now, the continuing war in Chechnya among other things, but the death of Alexandr Baev on 31 December last year has also forced a decision on the Russian contribution to the international Human Genome Project. Baev, who was 90 when he died, had previously been the biology secretary of the Soviet, then the Russian, Academy of Sciences before being put in charge of Russia's special programme in human genetics. In his earlier incarnation, Baev had been something of a conservative, but he appeared to have taken to his later role with the startled enthusiasm of a chauffeur put in charge of a country's air transport industry.

In truth, Russia's original contribution to the project has not been outstanding. Although formally a member of the unofficial Human Genome Organization, Russia has been too occupied with other things to do much more than keep up with what has happened elsewhere. Working scientists in Russian laboratories have access to the databanks, but more often from copies on magnetic tape than in real time. Yet Russia is not innocent of modern technology in this field. On the contrary, young people seem to have taken to the techniques of molecular biology with flair, sometimes in the knowledge that experience in the field may be a passport to a job elsewhere. None of this, of course, depends on the human genome project.

Baev's death would ordinarily not have been a setback, even though the funds for the Russian Human Genome Project are provided directly by the Ministry of Science. The ministry would have canvassed opinion and appointed a successor and the project would have continued. Baev's natural successor is Professor Andrei Mirzabekov, director of the Engelhardt Institute of Molecular Biology in Moscow, on whom Baev appears to have relied for advice. At Baev's funeral earlier this year, his succession seemed to be assured. But it will not be as simple as it seemed.

A lobby headed by the Novosibirsk academician Dmitry Knorre is now urging on the science minister, Boris Saltykov, that the Russian human genome project should be based at the Institute of Biological Chemistry at Novosibirsk. The argument advanced is that Mirzabekov is frequently overseas. That is so; on the strength of his proposal that the hybridization of random sequences of nucleotides can quickly tell the sequence of a single DNA strand, he now holds a joint appointment with the Argonne National Laboratory in the United States. But what the Novosibirsk cabal says is a disqualification is, in reality, the opposite. It will be some time before Russian science is able to take a commanding lead in this or any other comparable field. Meanwhile, participation in international ventures of this kind is a convenient way of transferring the technology of basic research. Why not give the task to the man best able to bring that about? □

Fusion scientists urged to support political campaigns

Washington. In a new and apparently desperate tactic to protect science funding from budget cuts, supporters of fusion research have been advised to enclose cash payments with letters to their congressmen seeking support for the troubled fusion programme.

The suggestion comes in a newsletter published by Fusion Power Associates, the main Washington interest group backing the \$360 million-a-year magnetic fusion programme. In an article headed "Fusion's last chance", the newsletter urges members of the organization to write to their member of Congress. "Also, you should think about sending a contribution to your congressman along with your letter," it says. "They appreciate support from their constituents, as it helps to compensate their campaign costs."

Congressional staff are sceptical whether such a tactic is likely to have a significant impact. Congressmen look to their districts primarily for votes, not money, says one, and take all personal letters from constituents seriously.

Also, it is illegal for congressmen to run their official business and their campaign fund-raising from the same office. But enforcement of this split is problematic. It may be illegal to record the fact that supporters of, say, fusion research have sent campaign money. But it is hard to believe that every congressman would loftily ignore a deluge of cheques from constituents interested in any single issue.

What frustrates friends of science in Congress, staff say, is the stubborn refusal of scientists as a group to engage in the political process at all, despite their almost total dependence on the public purse. Congressmen who devote great attention to science "never get a dime out of these folks," one staffer observed ruefully. Getting them to write letters, he added, "is like pulling teeth".

Steve Dean, president of Fusion Power Associates and author of the cash appeal, says the problem is that "scientists don't do this kind of thing and everyone else does". Suggesting sums of \$20 or \$50 as suitable for donation purposes, Dean says that scientists have already lost their long-cherished image as being aloof from the unseemly scramble for public funds.

Other science lobby groups are sceptical of the value of such a raw connection between cash and ideas. "We'd never do a thing like that," said the head of public affairs of a major scientific society. "It may be time for scientists to do it, but it isn't the time for us to tell them to do it." C.M.

Manhattan physicists cleared by FBI inquiry

Washington. Four nuclear physicists accused of spying for the Soviet Union in a book written by the former Soviet spy Pavel Sudoplatov, and published last year, have been cleared by an official inquiry carried out by the US Federal Bureau of Investigation.

The FBI inquiry found no evidence that J. Robert Oppenheimer, Enrico Fermi, Leo Szilard or Niels Bohr (all of them now dead) had knowingly supplied information to the Soviet Union while engaged on the Manhattan Project to develop the US atomic bomb. The allegations were made by Sudoplatov in his book *Special Tasks*, and were based on alleged contacts between the physicists and Soviet agents working in the United States.

The Clinton administration is expected to release this week the results of the FBI investigation, which was carried out at the request of the President's Foreign Intelligence Advisory Board (PFIAB). The inquiry was instigated by Sidney Drell, who is deputy director of the Stanford Linear Accelerator Center (SLAC) in California and a member of PFIAB.

Details of the exact wording of the announcement, as well as the official auspices under which it will be made, were still being agreed on Monday. But plans to issue a statement were confirmed by Randy Deitering, head of the PFIAB staff.

Sources close to the FBI's inquiry say that it found "no evidence whatsoever" that any

of the four physicists accused by Sudoplatov had "wittingly assisted the KGB" at any time. The original charges had caused widespread anger in the scientific community, and had led to a suggestion from the Federation of American Scientists that publishers should experiment with a form of peer review to avoid what it portrayed as an attempted character assassination of Oppenheimer and his fellow physicists.

The outcome of the FBI review was revealed by Drell to physicists attending a recent gathering at Cornell University held to celebrate Hans Bethe's sixtieth year at the university. Bethe, who won the Nobel prize in 1967 for work that he had carried out some 30 years earlier on the theory of nuclear reactions, was head of theoretical physics at Los Alamos at the time of the Manhattan Project.

Some physicists who attacked the Sudoplatov book when it was published at the beginning of last year claim that its allegations about atomic spies had already been thoroughly discredited, and privately question the value of the FBI probe. "It'll just give the book more publicity," says said.

But Drell insists that former colleagues and the families of the four physicists were entitled to see them officially exonerated, and says he is confident that this week's statement from the administration will accomplish that goal. **Colin Macilwain**

Letter from Einstein goes to auction

London. A 60-year-old letter (right) from Albert Einstein to Sir Karl Popper, who died last year aged 92, is expected to fetch between £6,000 and £8,000 when it goes on sale at Sotheby's in London next month.

In the letter, Einstein gives Popper a résumé of the celebrated imaginary experiment in which he casts doubt on the ability of quantum mechanics to give a complete description of physical reality. Einstein's proposed experiment would have invalidated one that Popper himself had suggested; Popper was said to have been "mortified", unaware that Einstein's views were not shared by many of his contemporaries.

The letter is part of a sale that includes Popper's complete 6,000-volume working library, which filled every room in his home — including outhouses, hallways and even the garage. Due to take place on 19 May, it will be the first auction in modern times

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

of a major philosopher's library.

The working library itself will be kept intact and sold as a single lot for an estimated £200,000. The master copies of Popper's books, including *The Open Society and its Enemies* and *The Logic of Scientific Discovery*, will also be auctioned as a single lot, and are expected to fetch £50,000. □

Congress turns spotlight on US drug approval agency

Washington. The US Food and Drug Administration (FDA) came under strong attack last week from both Republican and Democratic senators during hearings before the Congressional Committee on Labor and Human Resources. Most critics appeared to agree with the view of committee chairwoman Nancy Kassebaum (Republican, Kansas) that "everyone is in support of the FDA, but it could be significantly improved."

Criticisms of the FDA included charges of intimidation, claims that the agency's 'culture' is one of confrontation rather than cooperation, and that the agency — which is responsible for regulating drugs, medical devices, food and animal drugs — takes too long in approving products. But David Kessler, the FDA's commissioner said after the hearing that "some of the criticism may be of a past agency."

Senator Judd Gregg (Republican, New Hampshire) said the Congressional committee received daily complaints of unwritten rules in the agency. A letter to the agency from the House Committee on Commerce sent after a subcommittee hearing, and signed by leaders of both parties, claimed that during the hearing and in prior conversations "there appeared to be a widespread fear of retaliation by FDA personnel against witnesses, FDA employees and others who cooperate with the committee".

Barbara Mikulski (Democrat, Maryland) backed demands for reform. "There is enormous frustration with the FDA about management, enforcement and nitpicking and that must be acknowledged," she said. "We need a sense of urgency, and if we don't get it, Congress is going to roll right over you."

Other criticisms focused on complaints of the FDA's lengthy drug approval procedure. Committee member Bill Frist (Republican, Tennessee), who is a surgeon, complained of having to refuse treatment because a device available in Europe was not available in the United States. Kessler countered with an example of what happens when devices are not well regulated.

Kessler also disputed the charge that the agency is slow to approve new drugs, and pointed to reductions in average length of the review process over recent years as evidence of a general speeding-up of drug approval time. Mark Novitch, professor of health care at the George Washington University Medical Center, added that some of the reductions are a consequence of alterations in the way in which time is accounted for. But some reductions are real — and will be greater in future.

A representative of the biotechnology industry explained that the FDA's approval

time was only a fraction of a drug's total development time. George Rathman, president and CEO of the ICOS corporation and a pioneer of biotechnology, pointed out that the time needed to generate the data required for a new drug application is growing, now averaging almost seven years.

Asked by Senator Gregg whether he recognized the need for reform, Kessler replied that the FDA is a regulatory agency which sometimes has to say "no". Kessler added, however, that some changes were already being introduced following the release last week by the White House of a policy document on FDA reform.



Kessler: claims delays are being reduced

approval for their manufacturing processes on the basis of pilot rather than full-scale plant. The latter could save industry millions of dollars.

A spokesman for the Biotechnology Industry Organisation (BIO) said it welcomed the changes, but was disappointed that the proposals did not include outside reviews and deregulation of Phase 1 clinical trials, which examine drug action.

At present, proposals for such trials are submitted both to the FDA and to a local Institutional Review Board (IRB), made up of both scientists and lay people. The IRBs examine the design of studies, and ensure that human subjects are protected, while the FDA looks at the science involved. Industry and other observers argue that the Institutional Review Boards could also evaluate the science.

"We've had favourable reactions to this idea in conversations with people at the FDA, so we were disappointed not to see such a proposal," says Alan Goldhammer technical director of BIO. For now, proposals in the White House document requiring changes in guidelines or regulations will go ahead.

Proposals needing legislative changes must await Congress, which is likely to make some legislative changes of its own. Later this year, Kassebaum is planning to hold more hearings which will focus on specific reform proposals for the FDA.

Helen Gavaghan

London gets funds to restructure its medical colleges

London. Virginia Bottomley, the UK health secretary, has announced the government's approval of a set of new projects planned as part of the reorganization of London's hospitals around four multi-faculty colleges, each with high-class research facilities.

Bottomley was sharply criticized for appearing to ignore the opposition of groups that will see their local hospital facilities either reduced or eliminated completely as a result of the rationalization plans. But the moves have been broadly welcomed by London's biomedical research community.

The Higher Education Funding Council for England (HEFCE), for example, is to provide £20 million towards a new basic and medical science building at Imperial College. This will enable Imperial to cement a long-planned merger with the Royal Postgraduate Medical School in Hammersmith, and with the Charing Cross and Westminster Medical School. Sir Ronald Oxburgh, the rector of the college, last week described the plans as "the opportunity of a lifetime".

The funding council has also provisionally approved a grant of £14.5 million to University College to allow it to convert the nineteenth-century Cruciform Building, the original core of University College Hospital, into a new centre for teaching and research (the Wellcome Trust has agreed to provide an extra £11.5 million towards the centre).

The Department of Health itself is providing funding for a major neurosciences and neurosurgery centre at King's College, south of the Thames, the main hub of the third cluster. The decision to place the centre at King's — earlier reviews had proposed that it be situated at Guy's Hospital, also in the same cluster — partly reflects the growing reputation of the Institute of Psychiatry at the adjacent Maudsley Hospital.

The largest lump sum, an investment of £240 million, will go towards building new facilities for St Bartholomew's and the London Hospital. These are to be combined on the site of the latter in east London to form the core of the fourth cluster.

Under a fierce attack in the House of Commons last week over some of the implications of these moves — and in particular for using a press release to announce the planned closures — Bottomley insisted that they are an attempt at "concentrating centres of expertise so that they can compete into the next century".

But her remarks came too late to dissipate the heated public opposition to many of the closures such as that of the emergency services and hospital wards at St Bartholomew's' Smithfield site, where it has been active since the eleventh century (see *Nature*, 361, 194; 1993).

David Dickson

MITI drops headlights for gene sequencing

Tokyo. Japan's powerful Ministry of International Trade and Industry (MITI) has taken a further step into genome research by converting a former car headlamp-testing facility into a state-of-the-art centre for sequencing DNA.

At the end of last month, MITI officials and some of Japan's leading molecular biologists gathered to celebrate the completion of the new sequencing centre, which has been installed over the past two years on the fourth floor of the ministry's International Trade and Industry Inspection Institute in western Tokyo.

Until now, MITI's involvement in genome research has been limited, as the field has been considered the responsibility of other ministries and agencies. The Ministry of Education, Science and Culture (MESC) oversees university research in the area, and the Ministry of Health and Welfare (MHW) funds medical research and related areas. Similarly, the Ministry of Agriculture, Forestry and Fisheries has a rice genome project, and the Science and Technology Agency, along with the MESC and the MHW, funds part of the human genome project.

Anxious not to tread on others' toes,

Russian aid agency defends its actions

London. INTAS, the body set up by seventeen European states to channel funds to scientists working in Russia and other states of the former Soviet Union, has defended itself against charges of being excessively bureaucratic in the handling of grants.

The charges were made last month by the European Union's research commissioner, Edith Cresson, who said the European Commission was not prepared to provide increased funding to INTAS until it raised the effectiveness with which grants were managed (see *Nature*, 374, 203; 1995).

But in a statement issued in Brussels last week, INTAS says that, since it is handling public funds, it has had to ensure their "reliable and transparent management, distribution and transfer" in accordance with the legal frameworks of the newly independent states, its own member states, and the EU.

"The criticisms should be viewed in the context of these principles and the challenging, complex ever-changing environment of the NIS," the statement adds, claiming that "INTAS has succeeded in establishing adequate solutions to overcoming these obstacles," for example by negotiating favourable treatment for taxes and customs duties.

INTAS officials claim that 500 out of 1,000 multilateral research projects selected for support have received initial funding and are already operational. **David Dickson**

MITI's entry into genome research is under a facility entitled the Biotechnology Inspection Centre. The institute in which it is based used to be an important part of MITI's post-war drive to promote exports, but is undergoing a radical transformation as it seeks new tasks for its 460 technical staff.

Other floors at the institute will be devoted to inspection activities related to implementation of a new product liability law, and a law designed to prevent the proliferation of chemical weapons.

Work is already underway at the centre on sequencing the genome of *Pyrococcus shinkai*, a deep-sea bacterium collected from a volcanic vent in the Okinawa trough by Japan's deep-sea submersible Shinkai 6000. It is hoped that heat-resistant proteins from



the bacterium can be isolated for possible use in the biotechnology industry.

With technical advice from Hiroaki Shizuya of the California Institute of Technology in the United States, 26 staff at the centre have already completed a physical map of the bacterium's genome. Michio Oishi, of Tokyo University's Institute of Molecular and Cellular Biosciences and one of the centre's advisers, says the researchers hope to have sequenced the whole genome

within about two years.

The genome is quite short — just under 2 megabases in length — and its relatively low guanine-cytosine content makes sequencing comparatively easy. Nevertheless, Oishi claims that it will be a remarkable achievement if the centre is able to complete the sequencing by 1997. An international project to sequence the 12.8 megabases of the yeast genome, for example, is not expected to be completed until around the same time, while another to sequence the 4.7 megabases of the *Escherichia coli* genome is making even slower progress (see *Nature* 368, 383; 1994).

But many observers remain concerned about whether the researchers at the sequencing facility, in spite of its ambitious goals, will be able to interpret and apply the data it is collecting. The staff are enthusiastic about their work, and have access to the latest sequencing equipment and computers. But they have no experience in genome research.

Furthermore, MITI itself has only limited expertise to draw on from its own research institutes, while Japan's traditional bureaucratic hurdles make it hard for the ministry to involve university researchers.

Such expertise is not totally absent. MITI can use its new National Institute of Bioscience and Human Technology in Tsukuba science city, and the ministry is also backing another DNA sequencing facility, the Kazusa DNA Research Institute, funded by the Chiba prefectural government (see *Nature* 372, 6; 1994). Scientists at both institutes will be able to help in the interpretation of data produced by the new sequencing centre in Tokyo.

But some leaders of Japan's genome research are even sceptical about Kazusa, dismissing it as merely a "local project". Given such comments, MITI seems likely to face an uphill battle in winning domestic recognition for its efforts in the field.

David Swinbanks

Ecologist tipped for UK science post

London. Robert May, Royal Society research professor of zoology at the University of Oxford, is being tipped as the most likely successor to Sir William Stewart as the next chief scientific adviser to the British government, and head of the Office of Science and Technology.

May, 59, holds a PhD in theoretical physics from the University of Sydney, Australia, and was professor of biology at Princeton University in New Jersey before taking up his current post in Oxford in 1988. Although he has had little direct involvement in British science policy, he is widely respected as both a researcher —

for example for his work on population dynamics and the mathematics of ecosystems — and a research administrator.

No firm decision has been taken on the successor to Stewart, who took up his post in 1990, and whose contract comes to an end in June. Other names that have been mentioned as possible candidates include Alan Rudge, a director of BT and chairman of the Engineering and Physical Science Research Council, David Wallace, vice-chancellor of Loughborough University, and John Arbuthnott, vice-chancellor of the University of Stirling. **D.D.**

Meeting agrees on need for new targets for

Berlin. The nine-day-long climate conference in Berlin rescued itself from potential disaster last week by pulling off a last-minute compromise and agreeing that emissions of greenhouse gases should, in principle, be reduced after the year 2000.

In the closing session of the conference, signatories to the United Nations Climate Change Convention, drawn up in Rio de Janeiro in 1992, agreed to draw up a protocol setting targets for reductions of the gases, and a timetable for achieving such goals, within the next two years.

After the meeting, pressure groups and developing countries said they were disappointed that specific targets had not been agreed in Berlin. But they welcomed the conference's conclusions as a step forward that could lead to action in the near future.

The importance of the so-called Berlin mandate is that it admits that the convention, whose signatories agree to return greenhouse gas emissions to 1990 levels by the year 2000, does not go far enough. Scientists and pressure groups have long argued that this measure, which few countries are in any case on target to achieve, is inadequate to avert predicted climate disasters.

According to the International Panel on

Climate Change (IPCC), which assesses available scientific information and recommends response strategies (see below), scientific models suggest that a 60 to 80 per cent cut in emissions is necessary simply to stabilize current carbon dioxide levels in the atmosphere.

The major aim of the Berlin meeting, in the eyes of many signatories, was to reach agreement on how to reduce emissions after the year 2000. But it soon became bogged down in a stalemate between those countries that are heavily economically dependent on fossil fuels — such as the OPEC countries and the United States — and the small island states, which fear for their existence if global warming raises sea levels. The Alliance of Small Island States (AOSIS) had called for an immediate agreement to reduce carbon dioxide emissions by 20 per cent by the year 2005.

The conference also became polarized over the issue of 'joint implementation', a reference to projects designed to reduce global levels of greenhouse gases which are financed by one country but carried out in another.

The United States wants a credit system under which such an investment by an industrial country in a developing country — for

example, by providing clean energy technology or planting forests to act as carbon sinks — would count towards the former's efforts at reducing emissions.

But many developing countries see this as a way by which industrial countries could avoid domestic responsibility. They also argue that their own rate of development would be hindered by having to use fuel sources that are less efficient than the fossil fuel sources that the United States would be free to continue burning. As the major polluters, these countries argued, industrial countries should bear the full cost of lowering greenhouse emissions.

The issues seemed unlikely to be resolved. But the German Chancellor, Helmut Kohl, raised expectations when he opened the ministerial part of the meeting on 5 April with a speech reiterating Germany's commitment to a reduction in carbon dioxide emissions by 25 per cent by the year 2005.

Cynics pointed out that Germany has already succeeded in getting a considerable way towards this target already by closing down many of the inefficient and polluting factories in east Germany. "But we are still only half way to this target" says Stephan Singer from the German branch of the

Climate change panel to remain main source of advice

Berlin. The Intergovernmental Panel on Climate Change (IPCC), the body that collates data on global warming and analyses their social and economic implications, is to continue as the major scientific advisory body to governments that have signed the United Nations Climate Change Convention.

Last week, the so-called Conference of the Parties (COP) — made up of the 128 governments which have ratified the convention, drawn up in Rio de Janeiro in 1992 — agreed at their meeting in Berlin that the panel should be its main consultative body on scientific issues. It also resolved to provide 10 per cent of its planned budget of 4.0 million Swiss francs (US\$4.7 million) next year.

The rest of the money will be put up by the World Meteorological Office (WMO) and the United Nations Environmental Programme (UNEP), the two organizations that set up the IPCC in 1988 in an attempt to raise the awareness of politicians of the seriousness of climate change-related issues.

The panel, chaired since its inception by the Swedish climatologist Bert Bolin, has three working groups. The first looks solely at the scientific aspects of climate change; the second at the likely impact of such change — including the options for

mitigating its predicted effects — and the third at the economic implications, including an outline of future scenarios of energy use and social patterns.

The panel produced its first report in 1990, arguing for urgent steps to reduce the emission of greenhouse gases if the disastrous consequences of global warming are to be avoided. All three groups are working on a second full report. The report will be published in the autumn, having passed IPCC's complex refereeing system involving almost a thousand reviewers.

There is at least one representative from a developing country on the writing teams responsible for each of the report's 59 chapters. As with other IPCC documents, it will be presented to a plenary meeting, at which all signatory nations, as well as non-governmental groups, have a right to comment on its contents. The final document will be discussed by the next Conference of the Parties, due to be held in

1996, probably in Montevideo.

Bolin defends the operation of the complex IPCC system against charges that it inevitably mixes science and politics. He argues that its commitment to consensus — an approach which some fear can water down its scientific message — is essential for an issue as sensitive as global warming, whose economic and political consequences are so potentially significant. "It is not the same as considering how to handle radioactive waste, for example, whose consequences are relatively localized and easy to control," he says.

Bolin also argues that the continuing build-up of greenhouse gases in the atmosphere makes it important to reduce emissions as soon as possible. The economic and political difficulties of such a step make consensus particularly important, he argues, in order to prevent individual countries from exploiting any dissent. "This is political reality."

The signatory states to the convention appear to endorse this view. The 10 per cent of the budget to be provided by the COP tacitly "is just about the right level of contribution to allow us to retain our independence", says IPCC general secretary Narasimhan Sundararaman.

Last week's meeting in Berlin also agreed to set up two subsidiary bodies to



Bolin: consensus remains 'essential'

greenhouse gas emissions

World Wide Fund for Nature.

"Now that Kohl has put climate protection high on the political agenda it will be hard for the finance ministry to block the wishes of the environment ministry, as has happened in the past," says Singer. Germany will now have to look seriously at measures which will reduce emissions in the west of the country, he says.

Intense and prolonged negotiations followed Kohl's speech, continuing until the small hours on 7 April. The resulting agreement, known as the Berlin mandate, states that the present commitment under the Rio climate convention is inadequate, and establishes a procedure that will set voluntary targets for reduction of all greenhouse gases after the year 2000.

A working group will be set up to design the protocol for approval in 1997. The group will consider quantified targets for limiting or reducing emissions within a time-frame that has still to be decided. The agreement mentions the dates 2005, 2010 and 2020, but makes no specific recommendation.

In addition to the Berlin mandate, delegates agreed to launch a pilot phase for joint implementation projects. But, initially, the investing country will not be able to claim credit for reduced emissions in

the pilot phase. The situation will be reviewed in 1999.

Another decision agreed by the delegates was to set up subsidiary bodies to assess scientific data on global warming and to review the way in which countries are meeting their commitments (see below). They also decided to locate the convention's permanent secretariat in Bonn.

Germany's former capital, which will have plenty of office space when the government moves to Berlin in 1997, offered rent-free premises and support of up to DM3.5 million (US\$2.5 million) a year. The secretariat will receive from the United Nations further support of just under US\$9 million over the next two years, during which time it will build up a staff of 50.

Matthew Spencer, a spokesman for the environmentalist group Greenpeace, said after the meeting that many such groups regretted that the Berlin conference did not agree a target of a 20 per cent cut in greenhouse gas emissions, as had been demanded by many developing countries.

"But governments have eighteen months to find ways of doing this," he says. "So we are still optimistic, and we believe that there has been a tentative step forward."

Alison Abbott

the full Conference of Parties. One, the **Subsidiary Body for Scientific and Technical Advice (SBSTA)**, will convey scientific information to the COP.

The **Subsidiary Body for Implementation (SBI)**, will advise the COP on national implementation of the convention. Both subsidiary bodies will be open to delegates — who need not be technical experts — nominated by each of the 128 signatory states, and will hold their first sessions in October.

The SBSTA will be chaired by Tibor Farago, head of environmental policy in Hungary's Environment Ministry, and will be able to seek advice from any qualified scientific or technological advisory body — though the only body explicitly mentioned in this regard, is the IPCC. Some see a danger in this. "SBSTA will be composed of delegates and not scientists," says Arjet Stevens from Greenpeace International. "There is a risk that it could choose to take advice from bodies that are not genuinely independent."

Farago, a geophysicist who has worked as a climatologist, concedes that governments cannot be forced to delegate technically qualified people onto the SBSTA committee, but he hopes that the governments "will be motivated to do so". He describes IPCC as "one of the most important bodies" with which he will be cooperating.

Greenpeace and other environmental

groups would have preferred IPCC to remain the only official scientific advisory body. But IPCC says it does not feel threatened. "We are very happy with the proposal," says Bolin, who feels it appropriate that IPCC itself is not too closely tied to COP.

Scientists who have helped prepare the IPCC reports say they are pleased that its future as a key advisory body is now secure. "It is the only body that can do a good job," says Klaus Hasselmann from the Max Planck Institute for Meteorology in Hamburg. "It reflects the [main] view of the whole scientific community, and also conveys the degree of variability around that average view."

Bruce Callander of the United Kingdom's Meteorological Office, head of the technical support unit for IPCC's first working group, says that if the COP had decided not to use the IPCC, it would have had to reinvent the panel under a different name. "It has the goodwill of the best scientists," he says.

But both Hasselmann and Callander agree that this goodwill needs to be nurtured, and that scientists could rebel if asked to write a third report too soon. They point out that nothing is likely to happen in the near future to substantially change the IPCC's conclusions — and that the most important task for politicians now is to implement its recommendations.

A.A.



Yucca Mountain: long-term danger?

Academy may probe waste 'explosion' risk

Washington. The US National Academy of Sciences may be asked to assess controversial claims by two physicists at the Los Alamos National Laboratory that nuclear waste stored in a planned underground repository at Yucca Mountain, Nevada, could spontaneously explode.

The authors of a paper suggesting this possibility — Charles Bowman and Francesco Venneri — are both particle physicists and leading proponents of the use of particle accelerators for disposing of nuclear waste by transmuting its fissile isotopes.

Their paper says that plutonium and other fissile material stored in barrels underground would eventually leak, and that, mixed with rock that would act as a moderator, they could reach criticality and explode (see *Nature* 374, 204;1995).

Hazel O'Leary, the energy secretary, told the Senate Armed Services Committee last week that the academy had offered to assess the theory. Asked for her own assessment, O'Leary said she gave it "the credibility I would give to anyone who didn't have the assignment and, quite frankly, had another interest".

But Bowman and Venneri were defended by Richard Bryan (Democrat, Nevada), a strong opponent of plans to store civil nuclear waste under Yucca Mountain. "They are respected physicists, not people who got their degrees at some degree mill," Bryan said.

Bryan compared the two physicists to Galileo, who, he pointed out, had been branded a heretic in his day. He also alleged that Bowman had been muzzled by the management at Los Alamos — a charge that O'Leary promised to investigate.

Bowman and Venneri are now searching for a peer-reviewed journal prepared to publish their work. Jerry Saltzman, an official at the energy department's office of nuclear waste management, says that the department is waiting to see whether external peer reviewers support publication of the work, which was summarily dismissed by internal reviewers at Los Alamos itself. If so, the department plans to ask the academy to resolve the dispute.

Colin Macilwain

Money row looms over ESA role in space station project

Paris. Scaled-down plans for Europe's participation in the international space station, Alpha, announced only last month, have already become deadlocked after a dispute between member states of the European Space Agency (ESA) about how much each should contribute towards the overall costs.

Under the terms of the new plan, ESA has decided to reduce its proposed contribution from ECU3.5 billion (US\$4.5 billion) to ECU1.8 billion for the period from 1996 to 2000. This followed pressure from Germany and France, who both told ESA that they could not afford the original cost. But ESA has so far obtained commitments from its member states to cover just ECU 1.2 billion.

One problem is that Italy, which has serious budgetary pressures at home, has yet to make a firm commitment to the space station. Yet even if Italy does contribute ECU300 million, as originally planned — and many observers consider this now to be an over-optimistic figure — ESA will still be left with a deficit of ECU300 million.

France's national space agency CNES said this week it was not able to offer any more money. "We are not going to make up the deficit of others for a non-priority programme", says CNES chairman André Lebeau. "Germany leads the project, it should assume its responsibilities."

Jean-Daniel Lévi, director general of CNES, adds that the French space agency is not willing to let the space station "crush" its other programmes. In any case, no decision on the final level of the French contribution seems likely until after the presidential elections.

French feelings are running particularly high because the previously proposed Crew Rescue Vehicle (CRV) — which was to have been built by France, and would have cost ECU1.2 billion between 1996 and 2000 — is the main victim of ESA's scaled back plan.

"We wanted the CRV, but we didn't have the money for it", says Lévi. He still argues that ESA's proposals meet the cuts demanded by both Germany and France. But one consequence of the whole process is that France has reduced its planned contribution from ECU600 million to ECU300 million.

ESA's contribution to the station will now be limited to the Columbus Attached Pressurised Module (APM), which will be built by Germany and Italy, and the Automated Transfer Vehicle (ATV) a relatively low-price vehicle that will ferry freight and fuel to the station, and is to be built by France.

Observers say the CRV was more technologically challenging than both these contributions, and would have provided Europe with the expertise it lacks in returning astronauts to Earth.

The APM was retained, they say, because it was considered essential by the United States and Germany. The ATV is both cheaper than the CRV and also essential for the use of Europe's Ariane V launcher, intended as payment in kind as part of ESA's contribution towards the estimated ECU2.4 billion running costs of the station.

Germany seems to have different ideas. Speaking in a parliamentary debate on the national budget last week, Jürgen Rüttgers, minister for research and technology, said that the agency had not taken German demands seriously enough, and had not looked sufficiently hard for a solution to its financial problems.

"It's ESA's job to speak with the other partners and get this sorted out", says one official at the German science ministry. At present, he adds, France and Italy are "not giving responsible answers" to Germany's demands. Others point out that, given the high stakes involved, agreement will inevitably be reached, but it is likely to require a political input. **Declan Butler**

French research body wins back funding after budget crisis

Paris. The French government last week agreed to reimburse a FF500 million (US\$100 million) deficit accumulated over the past few years by the country's largest research organization, the Centre National de la Recherche Scientifique (CNRS). The sum represents the balance between promises of funds made in the budget and actual payments received.

But if the move has brought momentary relief to CNRS researchers, it has done little to dampen their concern about the government's long-term commitment to supporting research in general and CNRS in particular.

The deficit has dominated debate inside CNRS since last September, when Guy Aubert, the organization's newly appointed director general, told laboratories they could spend just 60 per cent of the funds allocated to them at the beginning of the year.

The subsequent row reached a climax in February, when the CNRS's national committee suspended a meeting with Aubert and passed a series of motions asserting that French research was being "sacrificed".

The researchers' protests seem to have paid off. Shortly afterwards, Aubert obtained a promise from François Fillon, the Minister for Research and Higher Education, that there would be no cuts or freezes in this year's CNRS budget of FF10.86 billion. He also promised to restore the FF500 million deficit, although on a schedule to be agreed after an official audit of CNRS's finances.

Last week's decision means that the government will unblock FF300 million immediately, and pay CNRS the remaining FF200 million before the end of the year. Fillon says he decided to take prompt action "in anticipation of the audit's results" in order to allay concern that the rescue plan might be affected by the outcome of the presidential elections in May.

But some observers suggest there may have been a more political motive at work. The government had rejected similar requests only a few days earlier, and the decision on reimbursement was eventually taken by Edouard Balladur — who is both prime minister and a presidential candidate — within hours of a public promise by Lionel Jospin, the Socialist candidate, to recapitalize CNRS fully if he was elected.

Some labour union representatives claim that the government's previous promises to help CNRS were intended primarily to quell protests from researchers — an allegation also made by Jospin. They remain concerned about the outcome of current discussions about a proposed reform of the agency, and by uncertainty caused by the forthcoming election. **D. B.**

Delphi poll backs English-French translation

Paris. French and German researchers agree that the United States heads the world science league, and that Japan is third; but both consider themselves to hold second place. This is one conclusion of a comparison of opinion polls of scientists carried out in both countries.

The preliminary results of the French poll were announced last week by François Fillon, the minister of higher education and research. Based on the Japanese 'Delphi' polling technique, 3,000 French scientists were asked for their views on 1,000 possible technological developments over the next 30 years. Scientists were asked to indicate how important they

thought each development would be, and when and how likely it was to be achieved. Among the priorities identified was a telephone link between France and the United Kingdom with simultaneous translation; optimistic that its gestation time would be less than that of the Channel Tunnel, the experts predicted such a link by around 2006.

The responses will be fully analysed by the middle of next year. Fillon says the results will indicate which areas deserve funding at a national level to preserve French strengths in the technology concerned, and which are candidates for international cooperation. **D.B.**

Error re-opens 'scientific' whaling debate

London. The leaking of a confidential paper questioning Norway's official count of minke whales in the North Atlantic has re-opened the debate about its methods of calculating whale stocks, and led to renewed calls from environmentalists for a halt to this year's planned whale hunt.

But the paper, which found three separate errors in the software used to estimate minke whales, seems unlikely to persuade the Norwegians to lower their estimates from the present figure of about 86,700 — or to stop the hunt from going ahead. One Norwegian researcher describes the new information, which suggests a lower figure of 60,000, as "relatively minor" and unlikely to have any impact on the country's whaling programme.

The leaked paper was written by Jon Helgeland and Gudmund Høst from the Norwegian Computing Centre and Tore Schweder from the University of Oslo, and was presented to a meeting of the International Whaling Commission (IWC)'s Scientific Committee in January.

The two scientists say that the main effect of correcting the errors they discovered has been to increase the scaling factor employed to allow for whales not seen during sighting surveys. Final whale estimates are arrived at by dividing whale sightings by this factor. A higher factor results in lower whale estimates; a re-run of the computer program with the errors corrected showed a doubling of the factor from previous estimates.

The paper, which was leaked to the press without the scientists' knowledge, is the sec-

ond such document to question Norway's estimates of the number of minke whales. Last year, Justin Cook of the World Conservation Union, a member of the IWC's Scientific Committee, was asked to validate Norway's whale estimates by running the Norwegian programme independently, came up with a figure of 53,000.

Norway has consistently refused to adhere to an IWC moratorium on hunting minke whales. Norwegian scientists say a total whale count in excess of 86,000 justifies a limited whaling programme — they have set a quota of 301 whales each season — and will not affect the overall whale population.

Greenpeace has demanded that the Norwegian government call off this year's whale hunt in the light of the new findings. "Norway's science is now exposed as a sham," says spokeswoman Alison Ross. "Norway has duped the IWC by claiming its whaling was justified by science. It must now cancel its hunt. The time for killing whales whilst fiddling the figures is over."

But Lars Walløe, scientific adviser on whaling to the Norwegian government, says that Norway has no intention of cancelling the hunt, which is scheduled to begin on 2 May — a month before the IWC's next annual meeting. Describing the software errors as "not grievous or serious", he says that the cumulative effect on whale estimates would be small.

Walløe also rejects suggestions of cracks in the Norwegian scientific consensus over whale estimates. He points out that the leaked document was not the only paper

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Minke whales: one fewer to count.

presented at the meeting.

According to Walløe, a figure of 87,000 minke whales is still on the low side, as calculations do not include recent "improvements" to the figures. Minke whales in the seas off the west coast of Scotland, in the Irish Sea and to the west of Ireland are not included in a recent sighting survey, he added. A 'worst-case scenario', he said, would put Norway's minke whale count at 70,000.

Ehsan Masood

Australia's CSIRO is urged to turn back to the future

Sydney. An internal report prepared by Australia's largest research organization, the Commonwealth Scientific and Industrial Research Organization (CSIRO), has recommended the abolition of a system introduced less than ten years ago, grouping its research divisions into separate 'institutes'.

A special committee appointed by CSIRO's board has also suggested that eight of its most senior positions be abolished, and that the organization return roughly to its old structure, under which each research division reported directly to central office.

The new structure would, it suggests, allow CSIRO to become more flexible and responsive to the needs of its customers — mainly industrial bodies. At the same time, rather than having a single senior officer at the head of the organization, the committee suggests it should be run by a small management team made up of a chief executive and five other senior executives.

The report of the committee, which included two board members, comes shortly after scathing criticism of CSIRO manage-

ment from a bipartisan Senate parliamentary committee. After reviewing evidence from scientists in the organization, the Senate committee concluded in a report published last December that the organization was over-managed and that research was being adversely affected (see *Nature* 372, 585; 1994).

But CSIRO's acting chief executive, Roy Green, denies that the latest report is connected with the findings of the parliamentary committee. He says CSIRO's internal committee was looking at the organization long before the Senate committee started work. The management committee, he adds, feels that the institute structure has worked well, even if it is no longer appropriate. "The committee's findings in no way reflect anything the Senate said," says Green.

The CSIRO institutes were set up in 1986, each constituting a grouping of divisions with broadly similar research areas, such as the Sydney-based Institute of Animal Production and Processing. They were intended to promote collaboration between

divisions with similar interests.

But according to the Senate committee at least, the plan has not worked out. On the evidence of its report, which strongly criticized the role of the institutes, their main function appears to have been to generate "a vast amount of paperwork without clear and worthwhile objectives".

Based on submissions from CSIRO scientists, the committee also described the organization's board as "ineffectual" and criticized the role of its chief executive, John Stocker. Stocker left the organization in March at the end of his five-year contract, and CSIRO is now searching for a new chief executive. The position has been advertised internationally, but the board has not indicated when it is likely to make a final choice to the full position.

Green says CSIRO is prepared to listen to comments on the committee's conclusions from its own scientists, but that it will be up to the incoming chief executive to decide which changes will be introduced.

Mark Lawson

Future of Japanese universities

SIR — I should like to comment on the article entitled "What future for Japanese universities?" (*Nature* 372, 711; 1994). Great expectation is being placed on Japanese universities to undertake serious self-reform. Since the end of the Second World War, Japan has experienced rapid and substantial changes, both socially and economically. These changes should, as a matter of course, have induced reform in both the content and approach of the universities' education and research programmes.

The greatest challenge facing Japanese universities is the enhancement of quality. A number of issues require urgent solution, including a lack of diversity in course curricula, a lax evaluation system and the resulting lack of competition, a closed and rigid system of institutional administration and a weak financial base.

Following guidelines issued by the University Council, individual universities have launched serious reforms aimed at creating institutional identity, improving educational quality, raising the level of education and research by expanding and reforming graduate programmes, and increasing the flexibility of higher education systems to accommodate the needs for adult and lifelong education.

But your article contains some misunderstandings about national universities in Japan. The first concerns the allocation of research funds. These are allocated in two ways, one based on the size of the recipient university, and the other on specific research proposals. Allocations based on a university's size provide the basic funding necessary to maintain its education and research operations, and are not allocated according to an evaluation of individual projects. It is up to the university to decide how these funds are distributed.

Chief among the allocations based on research proposals are the grants-in-aid for scientific research provided by the Ministry of Education, Science and Culture (Monbusho). These grants are awarded according to the independent judgement of the Committee on Scientific Research Grants within Monbusho's Science Council. The committee's review is assisted by some 2,000 university researchers, many of whom are nominated by the Science Council of Japan (a body representing Japanese scientists). The selection of research projects for funding is therefore made through strictly impartial and objective peer review. Apart from individual research grants, a similar allocation method for allocating funds is applied to support for large-scale science projects, based on thorough deliberation by the Science Council and its specialized subcommittees, and not arbitrarily by

officials of Monbusho.

Second, on the administrative autonomy of Japan's national universities, these are free to plan and implement, within the limits of their budgets, their own education and research programmes. Decisions are made at faculty meetings and by a university senate comprising representatives from all the faculties. The problem is that these faculty meetings and university senates do not necessarily function effectively. And, as pointed out in your article, there is no system for objectively evaluating university education and research activities carried out under conditions of administrative autonomy. Thus the system for allocating resources based on such an evaluation is not working particularly well.

We are fully aware of the need to solve these problems. But this cannot be automatically accomplished by "reconstituting the national universities as public corporations". The danger in viewing the administrative problems of Japanese universities in terms of their status as institutions is that we may fail to grasp the true nature of these problems. What is important now is to introduce the principle of competition and to establish an objective evaluation system in Japanese universities.

Katsuhide Kusahara
(Deputy Director-General,
Higher Education Bureau)
Ministry of Education,
Science and Culture,
3-2-2 Kasumigaseki,
Chiyoda-ku,
Tokyo, 100 Japan

Karl Popper

SIR — Sir Karl Popper was worthy of all the praise in the obituary (*Nature* 371, 478; 1994) by Hermann Bondi who, however, may have overestimated the living by stating that "the scientific community has lost the philosopher who many of us feel illuminated, far more than any other, the way in which we [ought to?] work".

In the preface to *The Logic of Scientific Discovery*, Popper states: "If we ignore what other people are thinking, or have thought in the past, then rational discussion must come to an end, though each of us may go on happily talking to himself." Bondi may be correct in supposing or knowing that investigators cannot ignore contemporary thought when trying to solve problems in branches of science, technology or medicine held in esteem today. Genetics, immunology and molecular biology have offered spectacular solutions to medical problems of our brave old species. It may be permissible to

forget thoughts in the past on problems solved or solvable by modern science but less so to overlook the fact that numerous multifactorial and other complex problems remain unsolved and might find solutions in our forefathers' thoughts.

Many problems drifted into fashion in the 1960s and 1970s and, though unsolved, they still belong in them, thus obscuring from our view promising approaches based on careful observations by physicists, chemists and physiologists decades or even centuries ago. Admirers of Popper have considered most disturbing conclusions about their problems based on the revival, by scientists "around the corner", of old approaches. Do many readily spot Popper's "black swan", the one that falsifies our conviction that (all) swans are white? Should authors be encouraged to define and put down potential falsifiers of their messages?

The above quotation was preceded by "... whenever we propose a solution to a problem, we ought to try as hard as we can to overthrow our solution, rather than defend it. Few of us, unfortunately, practise this precept; but other people, fortunately, will supply the criticism. ...". Are many of us able to provide a criticism based on thought in the past? Popper might have welcomed my [bracketed] addition to Bondi's first sentence.

Johan Ahlqvist
Sibbvik,
SF-25830 Västankfjärd,
Finland

Daedalus misses the boat

SIR — Daedalus¹ describes motion sickness as arising from the gut, despite well known overwhelming evidence to the contrary. Numerous studies (for example, ref. 2) involving humans and animal models have shown that the vestibular labyrinth is essential for motion sickness. In contrast, denervation of the gut does not prevent motion sickness.

Daedalus also states that motion sickness "occurs only in vehicles". In fact, motion sickness can be induced by making side-to-side head movements while spinning oneself in circles or by visual special effects when one is sitting stationary.

Finally, one has to wonder how well Daedalus expects to breathe after the DREADCO engineering servo-unit cancels all movement of the diaphragm.

Alan D. Miller
Rockefeller University,
1230 York Avenue,
New York,
New York 10021-6399, USA

1. Jones, D. *Nature* 373, 662 (1995).
2. Money, K. E. *Physiol. Rev.* 50, 1–39 (1970).

Cold War against the Vatican?

SIR — Is *Nature* waging war on the Vatican? It would seem so, at least judging from a recent leading article¹ and an even more recent Commentary by John Godfrey, "The Pope and the ontogeny of persons"². A reply by Helen Watt³ to this Commentary has already exposed some of its scientific weakness and misinterpretations. The Pope has correctly interpreted the facts of human reproduction and development in his statement that "... we should envisage human persons coming into being ... as each bright day is imperceptibly created during dawn"⁴.

We do not intend to argue on single scientific points, but to denounce Godfrey's falsification of data. To start with, he says that "[i]n his recent book *Crossing the Threshold of Hope*, Pope John Paul II airs his view on human reproduction". In fact the Pope discusses neither the biological nor the philosophical/theological aspects of human reproduction. He deals only with the abortion issue, from a social and moral standpoint, very briefly in one of the last chapters⁴. Again, Godfrey sounds persuasive in stating that "[u]ntil the late nineteenth century Christian tradition, following Aquinas, tended to grade the protection to the incipient person according to the stages of development". We would be curious to know which documents of the Christian tradition led Godfrey to this conclusion. Again, the truth is that from the very beginning of the Christian tradition the Fathers and Doctors of the Church, notably Didakè, Athenagoras, Tertullian and others, all the way to Pius IX, unanimously spoke against abortion at any stage of pregnancy. Even those who followed the opinion of delayed animation considered the act of abortion, if not a homicide, gravely immoral. Aquinas himself, despite his views on ensoulment, clearly dictated in those days by a virtually complete ignorance of the biology of human reproduction, condemned abortion from the moment of conception. Aquinas is also misinterpreted on his views about the concept of *Epikèia*, according to which there are exceptional situations when the law can be infringed for a good cause. But *Epikèia* refers to positive law, not to natural law.

Finally, Godfrey hints at the Pope's ignorance of modern genetics and embryology. Given that the Pope is not a scientist, as he himself admits², Godfrey must mean either that the Pope's scientific advisers are ignorant, or that he chooses

to ignore their advice. Godfrey will probably be surprised to know that possibly the clearest reflection of the Pope's view on the biology of human reproduction is contained in chapter 11, p. 65 of the Warnock Report⁵.

Angelo Serra

Giovanni Neri

*Istituto di Genetica Medica,
Università Cattolica,
Facoltà di Medicina "A. Gemelli",
Roma, Italy*

SIR — I congratulate Godfrey¹ on having understood in part the doctrine of the Roman Catholic Church. As he says, "the concept of 'person' is not only a marvelous theory; it is at the centre of the human ethos" (John Paul II: *Crossing the Threshold of Hope*). The origin of the human being is fertilization, the onset of the diploid phase. Haploid gametes (ovule and sperm) have never been considered men or women, persons or individuals having any rights.

Godfrey is right to say that fertilization is a long, complex process and that ontogeny is gradual. Rightly, there is no biological discontinuity that indicates when the soul is infused, but this is the main reason to protect the life of the fertilized egg: it may already be a human person and nobody can demonstrate the contrary. "*In dubio, pro reo.*" If someone does not respect human life from its very beginning and practises abortion, she or he is committing murder.

This is why Jesus Christ and His Church, the Roman Catholic, have defended human life from conception, that is to say, from fertilization, not from implantation or any later event.

J. A. Manzanera

*Universidad Politécnica de Madrid,
UD Anatomía, Fisiología y Genética,
28040 Madrid, Spain*

SIR — We can all appreciate Godfrey's efforts¹ to educate us in reference to contemporary thinking in the area of human embryology developments, but he, as others of similar mind, will have to continue to be disappointed if he expects to document a physiochemical event that announces the infusion of a soul into the human embryo. He presents interesting and well-known information regarding the cellular gradualism that characterizes mammalian fertilization, but this is hardly an appropriate argument that the 'humanization' of the embryo must also proceed as a continuum. One cannot expect the infusion of the human soul at conception to be a process that would lend itself to any form of measurement or physical documentation.

It is fascinating that Godfrey would

resurrect not only the philosophical but also the biological beliefs of St Thomas Aquinas and Aristotle to reinforce his own scientific/religious concepts. He further introduces the popular concept that the evolving organism can become human only within the framework of the developing central nervous system. Godfrey implicates myelination. Others have even employed the appearance of the primitive streak, an early phenomenon in central nervous system differentiation, to imply neurocellular uniqueness in the embryo.

Even the US National Institutes of Health Panel on Human Embryo Experimentation was willing to grant the embryo at all stages a uniqueness that sets it above other cellular systems. Yes, each human zygote in its own intrinsic biological form is constantly undergoing vigorous alteration and change based from the very beginning on its own genetic blueprint. In time, this self-contained molecular template can direct the evolving structure into becoming a completely formed and functioning human being.

Each of us is formed in the image and likeness of God. As a consequence, there will never be any scientific evidence that the human embryo at any stage of its development will demonstrate ensoulment.

Robert J. White

*Case Western Reserve University,
MetroHealth Medical Center,
2500 MetroHealth Drive,
Cleveland, Ohio 44109-1998, USA*

SIR — I am scared, thinking about the implications of Godfrey's declaration that "there is no moment when human life starts". If that is true, then personal (individual) responsibility loses its meaning, as we really do not know (or don't care) when this or that person started to exist. What would be the point in celebrating birthdays, for example? Does this mean that the whole of living humanity is just one living mass not unlike a bacterial colony or a tree (and not comprised of individual persons)? Without this respect for individual persons, I can see how Hitler could have easily used this declaration as a scientific justification for the Holocaust — he perceived the Jews as a corruption to the pure race, and disposed of them, much like a tree would shed leaves that have dried up.

Carlo A. Arcilla

*University of Illinois,
Department of Geological
Sciences (M/C 186),
845 West Taylor Street, Rm 2438,
Chicago, Illinois 60607-7059, USA*

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

1. *Nature* **371**, 185 (1995).

2. Godfrey, J. *Nature* **373**, 100 (1995).

3. Watt, H. *Nature* **373**, 379 (1995).

4. John Paul II *Crossing the Threshold of Hope* (Jonathan Cape, 1994).

5. *Report of the Committee of Inquiry into Fertilization and Embryology* (HMSO, London).

Virology institute changes direction

SIR — You reported recently (*Nature* 374, 299; 1995) my dismissal after almost eleven years as director of the Natural Environment Research Council (NERC)'s Institute of Virology and Environmental Microbiology (IVEM) in Oxford and less than three years before my retirement at 60. The dismissal was on "Compulsory Early Retirement (structure) grounds" although the institute will continue with a new director, that is, the same organizational "structure".

You also reported that NERC senior management says this move followed an independent review of IVEM work carried out in line with government policy as laid down in *Realising our Potential*. The only review of which I am aware was chaired by Professor Roger Whittenbury. It met once in London (at the end of December 1994). There were no oral or written presentations of IVEM work or outside peer review. The only scientific information concerning IVEM, and provided by NERC, was copies of the institute's annual reports (latest 1992–94) listing highlights of recent research, in particular work sponsored by industry, government and the European Union. Following this meeting, the group reported to NERC a month or so ago, recommending that a new working group be established to define NERC's priorities in environmental microbiology and thereafter the role of IVEM. The Whittenbury group did not make any detailed or substantive comments about the NERC research at the institute, or recommend a new mission or my dismissal. The most recent council mandated review of IVEM science and management (Science and Management Audit) involved a visiting group to the institute at the end of 1990. This was chaired by Professor Anne Warner, and its recommendations were implemented in full over the following 2–3 years. The next review was scheduled for the end of 1995.

NERC states that following a recent independent review, council decided to change the IVEM mission. Arguably, the new IVEM mission is wholly incorporated in the IVEM mission previously established by council and which the institute has implemented. I was told that "Council considers that this will require a new scientific leadership with capabilities for which you are unsuited".

There was no meeting of council to set aside the previous mission, to agree a new mission, to establish a working group, or to take a view about priorities in environmental microbiology, the IVEM science, the leadership required or my suitability. There was no notice or prior discussion

with me of either a new mission or my termination of service. No established procedure or due process has been followed.

David H. L. Bishop
8 Carey Close,
Oxford OX1 3SR, UK

Slow but steady

SIR — Although Daedalus usually describes new ventures likely to be successful, a recent proposal (*Nature* 373, 108; 1995) was partially applied long ago and proved unsuccessful. His forest factory contains tree trunks force-fed with a basal pressurized mineralized solution and branch terminals are alternately sucked to extract water and then pressurized with glucose solution to simulate circadian sap flow. He predicts "cheap and perfect factory timber" that grows "with amazing speed".

At the time when Great Britain obtained possession of the Cape of Good Hope in the early nineteenth century, many oak trees were necessary for the men-of-war of the Royal Navy, upon which the nation's power principally depended. In Britain oaks required three centuries or more to mature, and even then existing British forest supplies were recognized as needing replenishment. The Cape's climate was far more clement than Britain's and oaks were observed to grow there very rapidly, with projected ages at maturity of 80 or so. Oaks were planted in appropriate parts of Cape Town and surrounding regions, notably Stellenbosch, the second oldest settlement in the province. (This town still is known as Eikestad or the Town of Oaks). The oaks grew as rapidly as anticipated, so confirming Daedalus' deduction of the consequences of increased glucose production and mineral/water availability. Alas, the mature trees produced wood of grossly inferior quality, quite unsuitable for ships. But fortunately for the Royal Navy, steel had replaced oak in ship construction by that time. So advantages of controlled factory growth of trees probably do not include shorter times to maturity. Daedalus's forest factory should perhaps therefore be a long-term project, without force feeding.

S. Wynchank
Medical Research Council,
Parow, South Africa

Marcus Marci

SIR — This year we celebrate the 400th anniversary of the birth of Ioannes Marci Marci of Kronland. He is known to have been an outstanding scientist at University in Prague during the first half of the seventeenth century. His scientific activi-

ties covered, besides medicine and physics, mathematics, astronomy and philosophy, and 16 books were published in Prague, nine of them written during the Thirty Years' War.

Marcus Marci was born on 13 June 1595 in the town of Lanškroun in eastern Bohemia. After studying philosophy and theology at Olomouc Jesuit College he received the degree of philosopher in 1615, adding the title Boemus to his name to stress his Czech origin. In 1625, he was granted the degree of doctor of medicine, and subsequently received further honours. He was an official physicist of the Czech Kingdom and personal physician of the Emperor Ferdinand III and the Czech King Leopold I.

Marci wrote the first medical books written and published in Bohemia. In *De proportionem motus seu regula sphygmica* ("About a proportion of motion or the impact rule") (1639), he described the laws governing impact of spherical bodies. He published the results of his investigation of light properties in *Thaumantias. Liber de arcu coelesti deque colorum apparentium natura ortu et caussis* (see ref. 1). The results of his investigations marked a significant advance in physics² that was much later continued by Newton, Huygens and others. Already before 1639, Marci had realized that the period of a pendulum's oscillations is proportional to the square root of its length.

Marci had been working at the Prague University during an extremely difficult time. The Czech Protestant nobility was emigrating as a consequence of the defeat after the battle at White Hill (1620), a prelude to the horrors of the Thirty Years' War (1618–48). Despite these complicated times, Marci remained loyal to his Czech origins.

Since 1977 the "Ioannes Marcus Marci of Kronland Medal" has been awarded in recognition of achievements in spectroscopy. Marci is also among the names of scientists given to the craters on the Moon.

Miloslav Vobecký
Academy of Sciences of Czech Republic,
Institute of Analytical Chemistry,
Prague, Czech Republic

1. Marek, J. *Nature* 190, 1092 (1961).

2. Mach, E. *Die Mechanik in ihrer Entwicklung* 310–313 (Brockhaus, Leipzig, 1933).

Alive and shopping

SIR — "One of those to return from the war was the late Maurice Pryce. . ." says John Maddox (*Nature* 374, 211; 1995).

When I met Maurice Pryce dashing around Safeway the other day I couldn't help thinking he wouldn't often be late.

G. R. Walker
4714 W 7th Street,
Vancouver, Canada V6T 1C6

Astronomy under the Southern Cross

Australia's international standing in astronomy is increasingly built around two telescopes on a spectacular escarpment on the edge of the Great Australian desert.

Coonabarabran. The Southern skies are different, and much more spectacular. The Milky Way is more luxuriant than in the north; before midnight last week it conveniently ended in the distinctive bulge of the Galactic Centre just above the horizon. To northerners, the Large Magellanic Cloud, lying above the Galactic Centre, is startlingly made conspicuous by its unfamiliarity. So is the Southern Cross, and the dark patch (representing obscuring dust and gas) that it encloses.

All this is clear to the naked eye, without artificial aids except perhaps for spectacles. One night last week, an astronomer casually remarked that if astronomy had been founded in the Southern Hemisphere, it might not have taken as long to figure out that the Sun is but one of 100 billion stars in a disc-shaped galaxy. There are all the pieces of the Universe stretched across the sky!

But these are exceptional circumstances. The small township of Coonabarabran is 20 km away from the edge of the escarpment on which are perched more than half-a-dozen telescopes, among which the Anglo-Australian Telescope (AAT) is the most powerful. The others are the Schmidt Camera, first established by the British Science Research Council (now the Particle Physics and Astronomy Research Council), but now jointly owned with Australia. But the escarpment also houses the Siding Spring facility of the Mt Stromlo and Siding Spring Observatories of the Australian National University (ANU) at Canberra.

One remarkable feature of the escarpment is the way in which the rocks are studded with extinct volcanic cores, like Easter Island sculptures but on a truly gigantic scale. The altitude may be less than on Hawaii or in the Andes, and the weather less predictable, but the 'seeing' can occasionally be exceptionally good. Last week, observers were whooping with delight at the precision of the profile of a stellar image on a television screen, showing dispersion over only a fraction of an arc-second. "We do not always realize how good the seeing can be", said the director, Dr Russell Cannon.

The rest of us do not always realize what observers put up with — the spartan sleeping accommodation, the unfashionably early evening meals (so as to make the best use of the autumn night) and the semidarkness in which people work (so as to keep stray light to a minimum and to

make adaptation to the dark easily attainable). And then there are the long hours at the recording equipment and the frustration when a clear sky fills with cloud.

Otherwise, being an observer is often now no different from working in a laboratory anywhere. Although the AAT is equipped with a barrel-like cage into which an observer can clamber to spend the best part of a night, most people spend their telescope-time as data-gatherers, monitoring the recording equipment and only occasionally having to adjust the telescope so that the image of their star falls squarely on the slit of their spectrograph.

On the instrumental side, both telescopes are being used to push forward the technology of measuring simultaneously the spectra of several stars in the field of view. The incentive is the recognition that it is wasteful of the power of a telescope collecting light from some area of the sky to use it simply for recording the spectrum of the light from a single star, essentially throwing away the light from all the others visible. The idea is to place a single optical fibre exactly where the images of several stars should be. The fibres then feed the light from their respective stars to different points along the length of a spectrograph, yielding as many spectra as there are fibres. The technology is as fiddling as it could be; the ends of the fibres have to be dead-flat, and must be checked by interferometry. "It's different in communications", said one astronomer. "They can always amplify, but with us, if we lose a photon it's gone for good."

Last week, a woman graduate student from Lyons and a woman postdoctoral fellow from Australia were practising the craft (a little like embroidery, with all the fibres hanging loose) of cementing tiny prisms (at the end of optical fibres) to the surface of a glass plate in exactly the positions specified by an optical photograph of the patch of sky in question, conveniently displayed on a monitor. But the AAT is now building a machine to do the job automatically. There will be an array of 1,000 optical fibres, whose flattened ends will be positioned (in groups of 400) so to yield more than 1,000 spectra in a few hours of good seeing.

Even as things are, a Cambridge-based postdoctoral fellow (another woman) was observing last week that six of nine quasars turned up by such an exposure were apparently novel, not in the machine-

based catalogues to which such observations are referred.

People are delighted with the two telescopes, and with the 3.9-metre AAT especially. The pointing accuracy and stability are especially good, so much so that people are fearful of replacing the now obsolete computers with more modern equipment for fear of throwing a software spanner in the works. The latest project is to extend the field of view of the telescope to 2 degrees by fitting correcting lenses just before the prime focus of the telescope to correct for aberration at the edges of the present field of view. Again, the incentive is to wring as much information as possible from the light-collecting capacity of the telescope, but Cannon is also delighted that the AAT will be able to span the Southern Cross.

By its design, the 60-inch Schmidt camera is a different animal, originally placed here because of the need for a survey of the Southern skies. It has won its now-international reputation through the development of an emulsion (called '380') whose resolution is superior to those previously available. The benefits are obvious in the plates it yields, over a field of view seven degrees across. On the contact prints, it is as if a fine dust of spherical and elliptically shaped particles had been scattered on a sheet of clear plastic. Palomar is now said to be getting results as good in its second complete survey of the sky.

Between them, the two telescopes have given Australian optical astronomy a powerful impetus, even if radioastronomy remains its strong suit. Apart from the traditional bastions of Sydney and the ANU, other universities (such as Melbourne) are now building departments of astronomy. Opinion is divided about (but is mostly in favour of) an Australian application to join the European Southern Observatory, while Britain is no longer a reluctant partner in the two telescopes on this site. Indeed, the research council has just agreed that it will pay its half-share of the expenses for the coming five years.

None of this is a great surprise, given that the splendour of the Southern skies is a perpetual recruiting-sergeant for astronomy and all its works. It is noticeable that those in the trade, when well away from telescopes, are given to leaning their backs against the sides of their motor-cars and gazing upwards, as if to check that the Magellanic clouds are still there, and in their correct positions. **John Maddox**

Flight and phosphorylation

John C. Sparrow

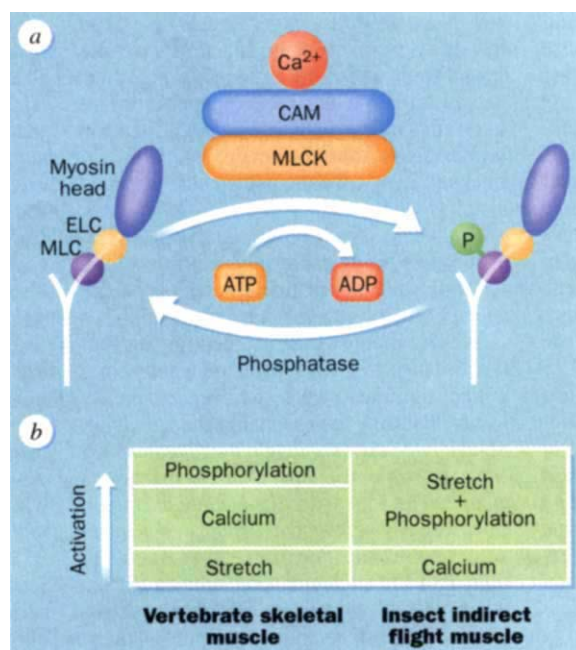
READERS of these pages will be familiar with the considerable impact of *Drosophila* genetics on our understanding of fundamental processes in development. But many research groups use *Drosophila* genetics to study other phenomena, and on page 650 of this issue¹ Tohtong *et al.* describe how they have employed a powerful mix of genetics and physiology to determine the importance of myosin phosphorylation for muscle function *in vivo* and *in vitro*.

Muscles are regulated by calcium release in response to nerve impulses. Several different mechanisms have evolved whereby the contractile apparatus responds to increases in intracellular calcium. In vertebrate skeletal muscle calcium activates contraction by binding to troponin-C, one of three troponins (I, T and C) which with tropomyosin form the regulatory complex of the muscle thin filament. However, the situation is not as clear-cut as this implies². Rapid and repeated stimulation of skeletal muscles to produce a tetanus increases phosphorylation of the myosin regulatory light chain (MLC) (*a* in the figure) and potentiates isometric twitch tension, apparently by increasing the calcium sensitivity of the contractile apparatus to activation. Because phosphorylation is faster than dephosphorylation, one advantage of repetitive muscle contractions, as in a warm-up immediately before an athletics event, is to enhance muscle performance.

The flight of insects such as *Drosophila* is powered by relatively large, stretched indirect flight muscles that are similar in structure to vertebrate skeletal muscle. But there are important differences between the two, largely because of the need of indirect flight muscle to power flight at wing beats of 200 Hz or higher. In flight, the muscle calcium levels are kept high by nervous input; the nerves fire at a lower frequency than, and asynchronously with, the wingbeat. The calcium is necessary, but not sufficient, to initiate flight, because full activation requires the muscles to be stretched as well³.

Stretch induces a delayed increase in tension. Opposing sets of indirect flight muscles shorten and are extended in an oscillatory fashion, the frequency being determined by the magnitude of the first-order delay between stretch and tension generation. This oscillation drives the

wings. The inertia of the wings, coupled to the elasticity of the thorax, makes a resonant system. Minimum power output from the indirect flight muscles is required when the frequency of muscle oscillation matches this resonant frequency; insects have evolved to provide this matching. The mechanism probably arose to minimize excessive calcium pumping — expensive in terms of energy and weight — which fixes 100 Hz as the probable limit of myoneural synchrony in insects⁴. Flight is



a, Phosphorylation of the myosin regulatory light chain (MLC) by the complex formed between calcium-activated calmodulin and myosin light chain kinase (CAM/MLCK). Only one of the two myosin heads is shown. ELC, essential light chain. *b*, Relative effects of calcium, stretch and light-chain phosphorylation in vertebrate skeletal muscle and insect indirect flight muscle. 'Activation' can be substituted by 'force', 'power' or 'myosin ATPase activity'.

initiated by myoneural calcium release, and a stretching of the indirect flight muscles when the initial jump of the fly deforms the thorax.

In *Drosophila*, the MLC is phosphorylated. The importance of phosphorylation in regulating flight-muscle activity was unknown until it was examined by Tohtong *et al.*¹, who mutated the MLC amino acids Ser 66 and Ser 67 (phosphorylatable serines corresponding to amino acids 15 or 16 in vertebrate MLC) to alanines and replaced the genes in flies lacking a functional copy of the single MLC (*Mlc2*) gene⁵. All the muscle myosin isoforms of these flies therefore contain MLCs which cannot be phosphorylated at these sites. Despite this, the flies develop into viable, fertile but flightless adult flies

(homozygous *Mlc2* null individuals are lethal). This is remarkable, showing at a stroke that no muscles that are important in fly development and survival have a requirement for phosphorylation of Ser 66 and Ser 67. The structure of the mutant indirect flight muscles appears to be normal, indicating that MLC phosphorylation is not required for myogenesis.

Tohtong and colleagues' mechanical experiments on skinned fibres (1–1.5 mm long) of indirect flight muscle show that phosphorylation is not required for calcium-activated force obtained at low strain (a smallish fraction of the maximal force obtained after strain); however, the absence of phosphorylation severely reduces the power output in the oscillating muscle which depends upon stretch activation. This demonstrates that there is a significant difference between vertebrate skeletal and indirect flight muscles (*b* in the figure). In the first, the major activation is calcium-dependent and is only slightly enhanced by MLC phosphorylation; the strain activation is small and transitory. In contrast, Tohtong *et al.* show for indirect flight muscles that full stretch-activation requires MLC phosphorylation and that its absence dramatically affects flight ability. In skeletal muscle, MLC phosphorylation causes a conformational change in the myosin^{6,7}, which may bring the myosin heads closer to their thin-filament binding sites². The relevance of this process to stretch-activation remains to be determined. The atomic structure of the myosin head with both light chains attached is known^{8,9}, so determination of the structural effects of MLC phosphorylation should now be possible.

The mesothoracic 'jump' muscles of flies are physiologically similar to vertebrate skeletal muscle³. The effects of calcium and phosphorylation on these muscles in wild-type and mutant flies may give insights into MLC phosphorylation in a predominantly calcium-activated muscle. Proposals² that MLC phosphorylation in skeletal muscle increases the calcium-sensitivity raises the

1. Tohtong, R. *et al.* *Nature* **374**, 650–653 (1995).
2. Sweeney, H. L., Bowman, B. F. & Stull, J. T. *Am. J. Physiol.* **264**, C1085–C1095 (1994).
3. Peckham, M., Molloy, J. E., Sparrow, J. C. & White, D. C. S. *J. Muscle Res. Cell Motil.* **11**, 203–215 (1990).
4. Smith, D. S. *Nature* **303**, 539–540 (1983).
5. Parker, V. P., Falkenthal, S. & Davidson, N. *Molec. cell. Biol.* **5**, 3058–3068 (1985).
6. Mrakovc-Zenic, A. & Reisler, E. *Biochemistry* **22**, 525–530 (1983).
7. Ritz-Gold, C. J., Cooke, R., Blumenthal, D. K. & Stull, J. T. *Biochem. biophys. Res. Commun.* **93**, 209–214 (1980).
8. Raymont, I. *et al.* *Science* **261**, 58–65 (1993).
9. Schröder, R. R. *et al.* *Nature* **364**, 171–174 (1993).

testable question as to whether this is its role in indirect flight muscle. In the wild, flies rarely have the opportunity for an 'athletic' pre-flight warm-up, because they require instantaneous maximum power. It seems likely either that significant MLC phosphorylation must occur within milliseconds or that the phosphorylation state will have to be maintained

at optimal levels during adult life — begging the question as to whether phosphorylation of the MLC has a regulatory or a modulatory role in insect indirect flight muscles. □

John C. Sparrow is in the Department of Biology, University of York, York YO1 5DD, UK.

GALACTIC CENTRE

The heart of the matter

Gerry K. Skinner

ACCORDING to Robert Frost¹

*We dance round in a ring and suppose,
But the Secret sits in the middle and knows*

but as we dance, we do get some hints about just what it is that sits in the centre of our Galaxy. Whatever is there is totally hidden from optical telescopes; interstellar dust clouds that lie in the plane of the Milky Way are thought to attenuate visible light by a factor of 10^{11} . Radio, infrared and X-ray telescopes can peer through the clouds. The trouble is that the stories they tell seem conflicting. Narayan *et al.*² may now have found a way of reconciling the existence of a massive black hole in the galactic nucleus, as implied by some observations, with the relative paucity of the high-energy emission that might be expected from such an object. On page 623 of this issue, they show how in some circumstances such a black hole can devour a surprisingly large fraction of the gravitational energy released by matter attracted to it.

Ironically, in many ways we know more about the inner regions of other galaxies than we do of our own galactic nucleus. A unified model³ is evolving which links a wide variety of active galactic nuclei (AGNs). At one extreme are highly luminous objects seen only in distant parts of the Universe and with appropriately exotic names (quasars, blazars and optically violent variables, to name but a few). At the other are 'Seyfert' galaxies, which are dominated by a bright point-like nucleus, and even galaxies that are normal except for an unusual source of energy at their core. All are seen as members of one extended family.

Many of the differences between the various types of AGN are explained as merely the effect of seeing similar objects from particular directions. Thus blazars are recognized as quasars in which one of the jets of relativistic particles is directed almost exactly towards us, boosting the apparent luminosity and giving the illusion of superluminal motion of the blobs of ejected particles. Similarly, two types of Seyferts are now widely thought to differ only in whether the orientation of a

dust-filled molecular cloud forming a torus around the nucleus is such as to allow us a direct view in or whether we see only scattered radiation.

Some AGNs have jets and associated radio lobes whereas others do not, but most of the remaining differences seem to be in the luminosity of the central 'engine' thought to be at the core of all these objects, which ranges from less than 10^{35} watts to more than 10^{40} watts. This central engine is generally thought to involve a massive black hole (MBH). The evidence is strongest in the case of quasars, where one can deduce that the emitting region must have a size less than 0.1 light years and a mass greater than 10^9 solar masses ($M_{\odot}$). Part of the potential energy released when matter approaches the event horizon of an MBH escapes as radiation and provides the power. Most calculations indicate that a fraction $q \approx 0.1$ of the rest mass energy mc^2 of accreted matter escapes in this way. An accretion rate of a few tens of $M_{\odot}$ per year, which can credibly be supplied by surrounding gas and by stars torn apart in the gravitational field, will provide $> 10^{40}$ watts and explain even the highest quasar luminosities.

Occam's razor and the economics of the automobile industry both suggest that where possible one should use the same components in the 'top of the range' models and in their less powerful brethren. So most models of Seyferts, and even of galaxies showing the slightest sign of activity in their nuclei, also tend to be based on accreting massive black holes.

It seems that in the past a higher fraction of galaxies were active and it has been suggested that many, perhaps most, galaxies go through phases where they show at least some nuclear activity. Once a galaxy develops an MBH at its core, it can never regain its virginity. So what of the core of our own, comparatively placid, Galaxy? Is there a central MBH engine, perhaps in a run-down or dormant state or in miniature form?

The motion of the stars and gas clouds immediately surrounding the nucleus of our Galaxy is most easily explained by supposing that they are moving in the

gravitational field around a central mass of about $10^6 M_{\odot}$ within a radius of 0.5 light years⁴, although no individual observations actually require such a mass. Remarkably, radio observations of the region reveal a compact, non-thermal, radio source termed Sgr A*, which is unique in our Galaxy. It lies exactly — within an arcsecond — where an MBH would be expected if one exists. Sgr A* is only about the size of the orbit of Jupiter⁵ and its (projected) velocity is less than 27 km s^{-1} , substantially lower than would be typical of an object of stellar mass moving in the gravitational fields of the region, so indicating a high mass (see ref. 6).

Thus there is circumstantial evidence that our galactic nucleus houses an MBH. The surprise is that Sgr A* is not a conspicuous object outside the radio waveband. Looking at infrared maps, the main centre of activity is the star group IRS16, an arcsecond away. There are still only hints of faint infrared emission⁷ from Sgr A*. And although there is variable X-ray emission from the region, presumably originating in Sgr A*, it is far weaker than expected⁸.

What Narayan *et al.* have now done is to produce a model for Sgr A* in which a black hole of mass $7 \times 10^5 M_{\odot}$ accretes matter but in which most of the energy is advected onto the MBH. The trick is that if the accretion rate is not too high the local cooling time can exceed the inflow time — the matter gets very hot but is swallowed before it has time to radiate much energy. With only a couple of free parameters to twiddle they predict a spectrum which matches that of Sgr A* remarkably well, from the radio band through to X- and γ -rays.

Of course, just because a model is consistent with observations it doesn't have to be the right one. With the Hubble Space Telescope and ground-based observations both finding evidence for large concentrated central masses in more and more galaxies, including comparatively nearby ones, there may be ample opportunity for testing the idea in a range of circumstances. □

Gerry K. Skinner is in the Space Research Group, School of Physics and Space Research, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

1. Frost, R. *The Secret Sits* (from *Complete Works*, Cape, London, 1951).
2. Narayan, R., Yi, I. & Mahadevan, R. *Nature* **374**, 623–625 (1995).
3. Antonucci, R. A. *Rev. Astr. Astrophys.* **31**, 473–521 (1993).
4. Genzel, R., Hollenbach, D. & Townes, C. H. *Rep. Prog. Phys.* **57**, 417–479 (1994).
5. Krichbaum, T. P. *et al. Astr. Astrophys.* **274**, L37–L40 (1993).
6. Blitz, L., Binney, J., Lo, K. Y., Bally, J. & Ho, P. T. P. *Nature* **361**, 417–424 (1993).
7. Liu, T., Becklin, E. E., Henry, J. P. & Simons, D. *Astr. J.* **106**, 1484–1489 (1993).
8. Predehl, P. & Trümper, J. *Astr. Astrophys.* **290**, L29–L32 (1994).

Origins of homochirality

Jeffrey L. Bada

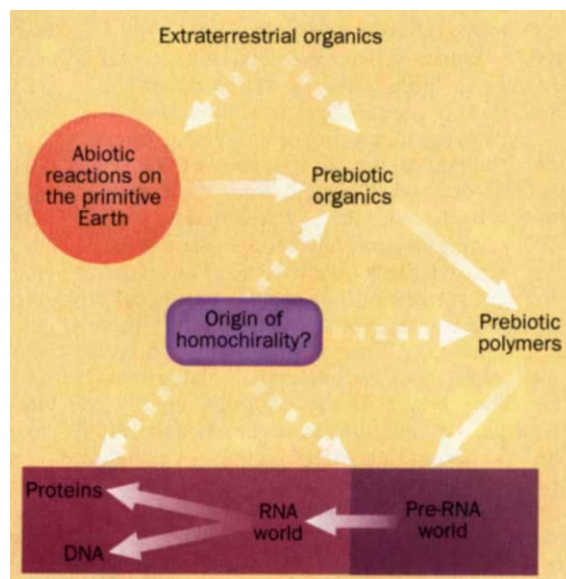
The process of mirror symmetry breaking is nowhere better illustrated than in the molecular components of living organisms. For instance, only L-amino acids are encoded into proteins and only D-sugars form the backbones of DNA and RNA, whereas laboratory syntheses of amino acids and sugars yield racemic mixtures (equal amounts of both the L- and D-enantiomers). A rigid stereochemical molecular architecture is present in all biomolecules that possess a chiral or asymmetric carbon. How and when did this homochiral characteristic of life originate — before, during or after the origin of life on Earth (see figure)? This question has intrigued scientists ever since Pasteur's discovery of the optical activity of biomolecules. As a recent conference* made clear, there is still no generally accepted explanation, but an emerging view is that biochemistry itself played a more important role than abiotic physical processes.

At the meeting, some of the attending physicists and theoretical chemists viewed homochirality as an inevitable consequence of universal fundamental physical processes which took place either directly on the early Earth or in extraterrestrial environments. Others considered a pure chiral medium to be necessary for the origin of life and that without molecular homochirality there could be no life. In contrast, some chemists considered biomolecular homochirality to be a consequence of life, rather than a prerequisite for life.

Physical mechanisms for the generation of biomolecular homochirality have been popular since the discovery of parity violation in weak interactions between elementary particles. How great a role could the spontaneous physical symmetry breaking in β -decay and weak neutral currents play in the origin of homochirality? The parity violation energy difference (PVED) between enantiomers is exceedingly small (around 10^{-17} kT) for most biomolecules. In the case of amino acids, the PVED would generate an excess of only 1 million molecules of the L-enantiomer in 6×10^{23} molecules, or 1 mole, of a racemate (R. Hegstrom, Wake Forest Univ.; A. MacDermott, Oxford Univ.; A. Mann, Univ. Pennsylvania; D. Cline, Univ. California, Los Angeles).

Amplification would be necessary for a

physical symmetry-breaking process to produce a significant enantiomeric enrichment. So autocatalysis is invoked, wherein a minute enantiomeric excess is greatly enriched. A nice example of autocatalysis is found in saturated solutions of sodium perchlorate, an achiral compound, which spontaneously crystallizes upon stirring into either pure left-handed or right-handed crystals (D. Kondepudi, Wake Forest Univ.). Could this occur under supposed geochemical conditions? In order for autocatalytic enantiomeric am-



When did homochirality begin?

plification to take place, extremely specific, and perhaps unlikely, conditions are required (Kondepudi).

Although homochiral origins through PVED and autocatalysis were suggested to have been possible on the early Earth, an extraterrestrial physical process involving the polarized synchrotronic radiation around neutron stars was championed by W. Bonner (Stanford Univ.) and M. Greenberg (Univ. Leiden). They suggested that in this astrophysical environment, a considerable enantiomeric enrichment in interstellar organic compounds could have been produced. In this case, exogenous delivery would then seed the Earth with a preordained molecular homochirality.

There is, indeed, a reported excess of L-alanine in the Murchison meteorite (M. H. Engel, S. A. Macko and J. A. Silfer, *Nature* **348**, 47–49; 1990). Is this evidence of an extraterrestrial origin of homochirality? In my view, it is dangerous to rely on the enantiomeric ratios of protein amino acids because of the omnipresent problem of terrestrial contamination. In fact, the

non-protein α -dialkyl amino acids in Murchison, such as isovaline, which are not prone to contamination problems, are racemic.

Any physical explanation for the origin of homochirality, regardless of where the process took place, must deal with spontaneous racemization (J. L. B.). This reaction would convert any enantiomeric enrichment back into a racemic mixture unless the production or delivery rates of the enantiomeric excess exceeded the rate of racemization. The racemization of amino acids on the modern Earth is a rapid geochemical reaction; it converts biological L-amino acids into racemic mixtures in 10^3 years at 50 °C and 10^6 years at 0 °C. Therefore, any homochiral amino acids on the primitive Earth must have been incorporated into important biomolecules and somehow protected from racemization on these timescales. Modern organisms still must deal with the problem of *in vivo* racemization and its detrimental effect on protein and enzyme structure and function, and this would also have been the case at the time of the origin of amino-acid homochirality.

The explanations based on fundamental physics were strongly challenged by chemists concerned with the molecules thought to be important in the origin and early evolution of life. This view was provocatively expressed in the title of one talk, 'Homochirality was not required for the origin of life' (S. Miller, Univ. California, San Diego). One conclusion was that D-ribose, the chiral sugar

present in RNA, is very unstable and decomposes so rapidly under geochemical conditions that it would have been an unlikely component of the prebiotic soup. Instead, perhaps the first self-replicating molecules present in the pre-RNA world were made up of achiral constituents (Miller). This theme was expanded upon in descriptions of the base-pairing and helical-forming properties of peptide nucleic acids (PNA) in which the backbone components are achiral (P. Nielsen, Univ. Copenhagen). Achiral PNA-like molecules are considered to be attractive candidates for the first carriers of genetic information because they form base-pairs and helical structures like DNA.

A further suggestion was that isotactic polymers (those containing monomeric units that are all the same handedness) might have formed randomly on the early Earth, although the number of D- and L-based polymers would have been identical (Miller). The importance of this can be seen with pyranose RNA molecules, where there is stereospecific base pairing

*The Origin of Homochirality in Life, Santa Monica, California, 15–17 February 1995.

even in a racemic mixture (A. Eschenmoser, ETH Zurich). According to this picture, one homochiral isotactic polymer acquired by chance a selective evolutionary advantage (for example, catalysis, or more efficient replication) and thus the origin of homochirality would have been closely associated with the origin of life. Miller emphasized that the origin of life was synonymous with the origin of evolution. It is possible that the homochirality of life is simply an evolutionary product.

Everyone agreed that the example of biomolecular homochirality on the Earth alone was inadequate to ascertain how widespread molecular mirror symmetry breaking is in the Universe. PVED-autocatalysis explanations predict that wherever life exists, the molecular handedness should be the same as on the Earth, whereas the biotic model predicts that life

throughout the Universe would be equally divided with respect to left and right molecular handedness. The resolution may come early next century, when we hope to explore Mars for evidence of extinct or extant life, and to analyse comets for organic compounds. The discovery of martian life based on D-amino acids would eliminate the PVED-based argument. Cometary samples will provide us with pristine material to determine whether there is a significant extraterrestrial enantiomeric enrichment in chiral molecules important in the origin of life. Perhaps these results will finally help to answer a question that has been asked for more than 150 years. □

Jeffrey L. Bada is at the Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093, USA.

PLANETARY SCIENCE

Buried dry ice on Mars

Robert M. Haberle

SINCE the 1970s, when the Mariner and Viking spacecraft gathered evidence that Mars might once have had a warmer and wetter climate, planetary scientists have been struggling to explain how it could have happened. The most popular theory is that the surface of Mars was warmed by the greenhouse effect of a much thicker carbon dioxide atmosphere than it has today. But if that is so — and there is some doubt¹ — where has the CO₂ gone? Was it stripped away by the last few giant impactors that smashed into the planet at the end of the accretion process? Did it evaporate into space by some exotic escape mechanism? Or was it irreversibly incorporated into the planet's rocks by weathering? Up until now, these have been considered the most likely reservoirs for the CO₂. But writing in the *Journal of Geophysical Research*², Bruce Jakosky, Brad Henderson and Mike Mellon examine another possibility: much of the original CO₂ may now reside as ice buried at depth in the polar regions.

The idea that the poles are reservoirs for CO₂ is not new, although the details are. In 1966, Robert Leighton and Bruce Murray³ suggested that the polar caps on Mars are carbon dioxide, and, because they survive at the poles all year long, their temperature controls the annual and globally averaged surface pressure through the equilibrium relationship with vapour pressure. The Viking spacecraft at least partly confirmed this prediction: the residual ice cap at the south pole is indeed CO₂. But in the north, the seasonal covering of carbon dioxide frost completely sublimates by the beginning of summer and exposes an underlying deposit of water

ice. This asymmetry in the behaviour of polar volatiles, combined with the rather low estimated density of the polar terrains (about 1 g cm⁻³) (ref. 4) and the more recent suggestion that the planet's obliquity varies chaotically⁵, motivated Jakosky and co-workers to examine the possible

buried beneath the water cap is CO₂ ice, or CO₂ in its hydrated form (CO₂·H₂O clathrate hydrate). For this to occur, the ice must be stable, which means it must be buried and it must be sealed off from the atmosphere. It must be buried, as its vapour pressure at the annually averaged polar surface temperature is greater than the surface pressure; and it must be sealed off from the atmosphere as it would evaporate otherwise. These conditions, Jakosky *et al.* argue, could be satisfied at a depth of only tens of metres. Furthermore, the total thickness of the ice could be as much as 4 km before basal melting would prevent a deeper deposit from forming. All told, Jakosky *et al.* estimate that the polar deposits could contain the equivalent of 0.85 bar of CO₂ as ice, or 0.2 bar of CO₂ as clathrate.

If true, the implications for the climate system on Mars are profound. As indicated above, buried CO₂ ice could be another reservoir for an early greenhouse atmosphere. But even more tellingly, it could mean that the present martian climate fluctuates wildly with the obliquity cycle. At high obliquity the water ice cap would sublime and expose the buried CO₂ to the atmosphere, allowing it too to sublime. As atmospheric CO₂ increased, so too would heat transport into the polar regions, and the system would run away until all the CO₂ was in the atmosphere⁶. The climate on Mars would be transformed into a warmer, more clement regime. The extent of the transformation would depend on many factors, but it would definitely be a different place from the one it is today.

Is there any evidence that Mars has seen substantial climate change in the relatively recent geological past? Yes, there is. Both polar regions are characterized by layered sedimentary deposits that many planetary scientists have attributed to quasi-periodic climate change associated with variations in the planet's orbital parameters⁷. But these variations had been thought to change the size of the atmosphere by factors of two or three at times of high obliquity, not the factors of 30 to 100 suggested by Jakosky *et al.*². And more recently, it has been suggested that various landforms on Mars are the result of a CO₂-greenhouse warming associated with volcano-driven episodic groundwater flooding of the northern lowlands⁸. This theory is controversial, but the greenhouse warming it requires could be

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Mars during the northern winter, with the ice cap in white.

composition of the polar deposits.

The present obliquity of Mars is about 25°, but could have varied from near 0° to as high as 60° during the past 10⁷ years. At the highest obliquities the summertime insolation at the poles would be much greater than it is today, so Jakosky *et al.* calculated how fast a water ice cap would sublime under such conditions. They found that even a water ice cap that was several kilometres thick would not last for much of the 10⁵-year obliquity cycle.

They then explored the possibility that

Sally Benson (1982)/SPL

1. Kasting, J. F. *Icarus* **94**, 1–13 (1991).
2. Jakosky, B. M., Henderson, B. G. & Mellon, M. T. *J. geophys. Res.* **100**, 1579–1584 (1995).
3. Leighton, R. B. & Murray, B. C. *Science* **153**, 136–144 (1966).
4. Malin, M. C. *Geophys. Res. Lett.* **13**, 444–447 (1986).
5. Laskar, J. & Robutel, P. *Nature* **361**, 608–612 (1993).
6. Haberle, R. M., Tyler, D., McKay, C. P. & Davis, W. L. *Icarus* **109**, 102–120 (1994).
7. Kieffer, H. H. & Zent, A. P. in *Mars* (eds Kieffer, H. H. *et al.*) 1180–1218 (Univ. Arizona Press, 1992).
8. Baker, V. R. *et al.* *Nature* **352**, 589–594 (1991).

produced by the mechanism favoured by Jakosky *et al.* rather than groundwater discharge.

So what can one conclude from all this? The available observations cannot tell us if the material underlying the residual north cap is water ice or dry ice. Nor can they tell us if oceans were present during the past several billion years, or if Mars did have a substantial CO₂ atmosphere. What they

do tell us is that Mars was different in the past, and that somehow water was able to flow across its surface. But until we return to Mars and take the next step in the exploration of this fascinating planet, all we can do is speculate. □

Robert M. Haberle is in the Space Science Division, NASA Ames Research Center, Moffett Field, California 94035, USA.

ENZYME MEMORY

What is remembered and why?

Alexander M. Klibanov

MOLECULAR memory is the curious phenomenon whereby functional properties of a substance depend on the sample's history. For example, when methacrylic acid and ethylene glycol dimethacrylate are copolymerized in the presence of the bronchodilator theophylline and then the latter is washed out, the resultant cross-linked polymer is capable of binding this drug; the same polymer prepared in the absence of theophylline under otherwise identical conditions cannot¹. Ever since the first report of enzyme memory in organic solvents², questions have persisted about its mechanism. Now, X-ray crystallography by Yennawar *et al.*³ has shed some light on this issue.

The original observation² was that lyophilization of the protease subtilisin from aqueous solution containing various competitive inhibitors (followed by their removal by anhydrous extraction) created an enzyme which was up to 100 times more active in anhydrous solvents than the enzyme lyophilized in the absence of ligands. This ligand-induced enzyme memory extended to the stability, affinity and substrate specificity of subtilisin in organic solvents, but it vanished when the enzyme was redissolved in water. Ståhl *et al.*⁴ subsequently demonstrated that the enantioselectivity of another protease, chymotrypsin, in organic solvents was also markedly affected by the chirality of the ligand present in aqueous solution of the enzyme during dehydration.

To rationalize these findings, it was hypothesized^{2,4} that the ligands caused conformational changes in the enzyme molecules and that even after removal of the ligand such 'imprints' were retained by the enzyme in anhydrous media because of its rigidity in the absence of water⁵. Because the structures of ligand-imprinted enzymes are distinct from those of their non-imprinted counterparts, so are the catalytic properties.

To assess this, Yennawar *et al.*³ soaked crystals of chymotrypsin — the same enzyme studied by Ståhl *et al.* — in hexane containing the enzyme inhibitor *N*-acetyl-D-tryptophan. Although the inhibitor

bound to the active site of chymotrypsin in this solvent, they found that the X-ray crystal structure of the enzyme in the complex was essentially identical to that of the free chymotrypsin in hexane⁶.

The very fact that enzyme studies in aqueous solutions are reproducible — that given characteristics (activity, pH-profile, stability and so on) of an enzyme are independent of where, when and by whom the measurements have been conducted — indicates that the behaviour of enzymes in water is invariant of their history. This lack of memory stems from conformational flexibility of enzymes in aqueous solution^{7,8}, where they are in a state of intensive perpetual motion or 'breathing'. Whatever reversible conformational changes, including ligand imprinting, may have occurred in the enzyme molecule beforehand are irrelevant, because in water the enzyme will shake them off and acquire the thermodynamically favoured conformation.

Thus one condition for enzyme memory, which is not met in water, is high conformational rigidity in the medium where the memory is to be displayed. The study by Yennawar *et al.*³ points to a further requirement — considerable conformational flexibility in the medium where the enzyme memory is induced by a ligand — which is met in water but not in hexane. On the other hand, both conditions are satisfied when an enzyme is imprinted with a ligand in water, followed by lyophilization and ligand removal, and then the memory is tested in anhydrous solvents².

There is growing spectroscopic evidence that on lyophilization from aqueous solutions many proteins undergo profound, reversible, conformational changes^{9–12}. This denaturing can be prevented to varying degrees, however, by adding certain excipients (lyoprotectants) to protein solutions before lyophilization^{11,12}. These findings have expanded the range of chemical messages that enzymes remember in anhydrous solvents beyond active site-binding ligands^{2–4} to include certain sugars¹³ and salts¹⁴ that bind to other parts of the enzyme molecule.

Enzymatic conversions in non-aqueous media broaden the utility of enzymes as practical catalysts, in particular for the synthesis of chiral pharmaceutical intermediates¹⁵. Imprinted enzymes hold a special promise because of their superior catalytic activity^{2,13,14} and even adjustable enantioselectivity⁴ in organic solvents. To take full advantage of this unique opportunity, much more must be learned about the strength and scope of enzyme memory, as well as the relationship between the nature of the imprinting excipient and the ensuing function of the enzyme preparation. Such knowledge can only come from structural investigations.

If enzyme crystals are to be used¹⁶ in organic solvents, the work of Yennawar *et al.*³ shows that X-ray crystallographic studies should be instructive provided that the crystals are grown in water in the presence of a ligand, followed by replacement of water with an organic solvent containing the same ligand. Limitations of this strategy are the likely heterogeneity of the resultant enzyme populations and difficulties in crystallizing them.

Neither these limitations nor restrictions on the type of memory-inducing excipients exist with lyophilized enzymes in organic solvents^{2,4,13,14}. Here such biophysical techniques as Raman and Fourier-transform infrared spectroscopies^{10,11} can be informative, although additional methodologies for the structural analysis of amorphous protein samples are certainly needed. □

Alexander M. Klibanov is in the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

1. Vlatakis, G., Andersson, L. I., Muller, R. & Mosbach, K. *Nature* **361**, 645–647 (1993).
2. Russell, A. J. & Klibanov, A. M. *J. biol. Chem.* **263**, 11624–11626 (1988).
3. Yennawar, H. P., Yennawar, N. H. & Farber, G. K. *J. Am. chem. Soc.* **117**, 577–585 (1995).
4. Ståhl, M. *et al.* *J. Am. chem. Soc.* **113**, 9366–9368 (1991).
5. Burke, P. A., Griffin, R. G. & Klibanov, A. M. *Biotechnol. Bioengng* **42**, 87–94 (1993).
6. Yennawar, N. H., Yennawar, H. P. & Farber, G. K. *Biochemistry* **33**, 7326–7336 (1994).
7. Karplus, M. & McCammon, J. A. *Rev. Biochem.* **52**, 263–300 (1983).
8. Rupley, J. A., Gratton, E. & Careri, G. *Trends biochem. Sci.* **8**, 18–22 (1983).

9. Poole, P. L. & Finney, J. L. *Biopolymers* **22**, 255–260 (1983).
10. Prestrelski, S. J., Arakawa, T. & Carpenter, J. F. *Arch. Biochem. Biophys.* **303**, 465–473 (1993).
11. Prestrelski, S. J. *et al.* *Biophys. J.* **65**, 661–671 (1993).
12. Desai, U. R., Osterhout, J. J. & Klibanov, A. M. *J. Am. chem. Soc.* **116**, 9420–9422 (1994).
13. Dabulis, K. & Klibanov, A. M. *Biotechnol. Bioengng* **41**, 566–571 (1993).
14. Khmel'nitsky, Y. L., Welch, S. H., Clark, D. S. & Dordick, J. S. *J. Am. chem. Soc.* **116**, 2647–2648 (1994).
15. Crout, D. H. G., Roberts, S. M. & Jones, J. B. (eds) *Tetrahedron: Asymmetry* **4** (spec. issue), 757–1080 (1993).
16. St Clair, N. L. & Navia, M. A. *J. Am. chem. Soc.* **114**, 7314–7316 (1992).

Bicarbonate briefly CO₂-free

Roger C. Thomas

ANIMAL life is a remarkably complex mechanism for converting oxygen to carbon dioxide. The subsequent reaction of CO₂ with water to make bicarbonate and H⁺ gives bicarbonate a central role in pH homeostasis. Numerous different membrane-transport mechanisms for bicarbonate are known, and another one

emerges from the ingenious studies with out-of-equilibrium solutions reported by Zhao and colleagues on page 636 of this issue¹.

The slowness of the reaction of CO₂ with water is overcome in physiology by the enzyme carbonic anhydrase^{2,3}. This was famously first found in red blood cells

where it maximizes the capacity of blood to carry CO₂ (as bicarbonate). The first bicarbonate transporter recognized was indeed the chloride/bicarbonate exchanger of the same cell. It allows the rapid interchange of bicarbonate and chloride, to generate the chloride shift. Since then many other bicarbonate transporters have been discovered, the one of most general importance in intracellular pH regulation probably being the Na-dependent Cl⁻/HCO₃⁻ exchanger^{4,5}.

A similar mechanism regulates pH in

OBITUARY

Joseph Needham (1900–95)

JOSEPH Needham died peacefully at his home in Cambridge, UK, on 24 March, at the age of 94. For the last few years he had been suffering from increasing loss of peripheral motor control because of Parkinson's disease; his end was hastened by a mild stroke on the evening of 23 March. True to his invariable habit he had spent much of that day at his desk in the Needham Research Institute, where he worked on letters and papers with the help of his assistant.

Needham will be remembered for his massive achievement embodied in the *Science and Civilisation in China* project, the successive parts of which have been published by Cambridge University Press since 1954. This great work is planned as a history of science, technology and medicine in China, seen in its fullest social and intellectual context, and illuminated by a deep and sympathetic understanding of the cultures of both East and West. Through his writings Needham has radically changed the ways in which scholars and scientists evaluate both the history of Chinese culture, and the history of science, medicine and technology understood as part of the common cultural heritage of the human race. He was undoubtedly the greatest Western sinologist of this century.

Needham was born on 9 December 1900, as the only son of a Harley Street physician and a musically talented mother. After attending Oundle School he went up to Gonville and Caius College, Cambridge, and read biochemistry. Caius was to remain his academic home for the rest of his life; he was successively a research fellow, tutor, fellow and finally (1966–76) Master. For most of the first half of his life Needham was engaged in establishing himself as a chemical embryologist of distinction. The major works of this period are his *Chemical Embryology* (1931) and *Biology and Morphogenesis* (1942). But by the time this second book appeared he was already moving in the direction which was to lead him towards his life's work.

In the mid-1930s he met three Chinese researchers who had come to work in Cambridge. The interest these bright young people aroused moved him to begin learning Chinese, and when war broke out in Europe and the East it was this connection that led him to propose that he should be commissioned to establish a Sino-British science co-operation office in Chongqing, to where

return to Cambridge he had already planned the years of work that lay ahead. He set out to answer a question that had been presenting itself to him ever more clearly for some time: why was it that despite the immense achievements of traditional China it had been in Europe and not in China that the scientific and industrial revolutions occurred? He approached Cambridge University Press with a proposal for a one-volume treatment of this subject, which they accepted. But as time went by this plan swelled to seven volumes, the fourth of which had to be split into three parts — and so it went on, the fifth volume being at eight parts and still growing. Sixteen parts in all have been published so far, and about a dozen more are on the way. Most of the earlier volumes were written in their entirety by Needham himself, but as time went by he gathered an international team of collaborators, to whom the completion of the project is now entrusted. As the project has broadened, so has the range of questions under investigation. It is now clear that no simple answer to Needham's original question will be possible. The quest has opened out into an investigation of the ways in which scientific and technical activity have been linked with the development of Chinese society over the past four millennia.

The home of the project is now the Needham Research Institute in Cambridge, which also houses a unique library used by scholars from both East and West. Much of the funding for the Institute's building came from donations from Hong Kong and the United States. Joseph Needham is dead, but through the vigorous intellectual life of the Institute, his work continues.

Christopher Cullen

Christopher Cullen is in the Department of History, School of Oriental and African Studies, Thornhaugh Street, Russell Square, London WC1H 0XG, UK, and chairman of the publications board of the Needham Research Institute.

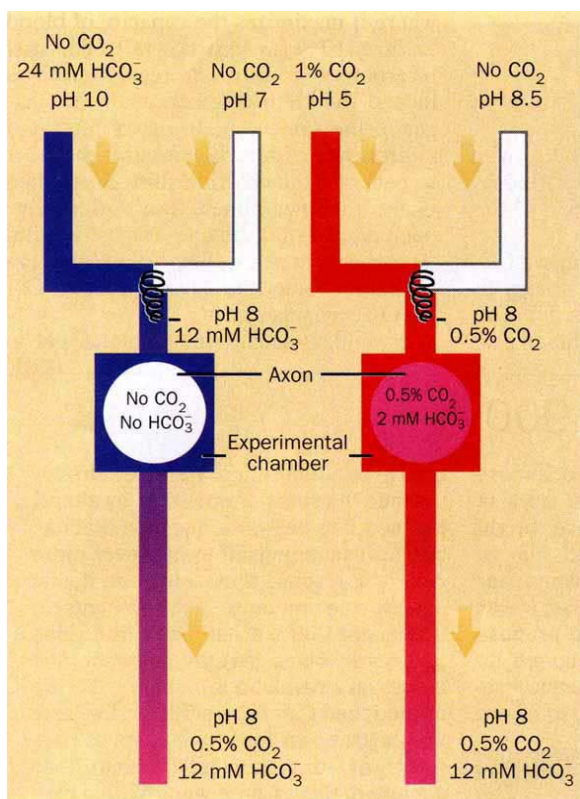
IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Joseph Needham — from embryology to sinology.

the Chinese government had withdrawn in the face of the Japanese onslaught. During this time he was ideally placed to study what had been accomplished by the Chinese people in science and technology over their long history. What he began to learn astonished him. It became clear for instance that printing, the magnetic compass and gunpowder weapons were all Chinese in origin, despite the puzzlement that Francis Bacon had expressed over their beginnings when in the seventeenth century he pointed to "the force and virtue and consequences of discoveries" (*Novum Organon*, Book 1, aphorism 129).

After the war Needham worked with Unesco in Paris for a while, but on his

Needham Research Institute



The continuous-flow apparatus and technique employed by Zhao *et al.*¹ to discover the K/HCO_3^- cotransporter. The significance of the work lies not so much in the cotransporter, the physiological function of which is not known, as in the imaginative experimental protocol involving out-of-equilibrium solutions.

squid giant axons⁶ but bicarbonate is still carried across the axon membrane in the complete absence of Na^+ or Cl^- ions¹. It proved difficult to investigate the mechanism involved using conventional CO_2/HCO_3^- solutions, because the CO_2 makes it impossible to maintain a significant bicarbonate gradient. The normal bicarbonate saline of pH 8 used in squid-axon experiments contains 12 mM bicarbonate and must be equilibrated with 0.5% CO_2 to be stable. If this equilibrium is disturbed, for example by the loss of CO_2 to the atmosphere, the pH will slowly increase as bicarbonate too is converted to CO_2 and lost.

A stable bicarbonate solution must therefore contain CO_2 . This crosses cell membranes so fast that its level inside cells is only rarely different from that outside. The intracellular CO_2 is in equilibrium with intracellular H^+ and HCO_3^- , often catalysed by the almost ubiquitous carbonic anhydrase. So you cannot normally have external bicarbonate without a similar level inside cells. The exact level will depend on the intracellular pH, but this is usually close to the external pH.

Zhao and colleagues¹ solved the problem by the clever use of out-of-equilibrium solutions (see figure). They mixed two pairs of carefully formulated solutions in a continuous-flow apparatus

which delivered the mixture to the outside of a squid axon within half a second. They mixed either a 24 mM bicarbonate alkaline solution or a 1% CO_2 acid solution with acid or alkaline CO_2 /bicarbonate-free solutions to give pH 8 mixtures with 12 mM bicarbonate or 0.5% CO_2 . The rate constants of the conversion of bicarbonate to CO_2 and vice versa are such that very little takes place within half a second. All solutions contained acetazolamide to inhibit any carbonic anhydrase that might be present.

That the CO_2 -free, freshly mixed solution indeed contained almost no CO_2 is shown convincingly by its lack of effect on the intracellular buffering power, which Zhao *et al.* measured by applying ammonium solutions. The equilibrated CO_2/HCO_3^- solution increased buffering power by 57%, as expected, whereas the out-of-equilibrium solution increased it by only 7%.

In contrast to its lack of effect on buffering, the CO_2 -free, out-of-equilibrium

bicarbonate solution doubled bicarbonate uptake rates by potassium-depleted axons in high-K solution, strong evidence for cotransport of K and HCO_3^- . Similarly, bicarbonate efflux from K-loaded axons in bicarbonate-free solutions was much faster if the bathing solution contained CO_2 . The CO_2 outside is converted to about 2 mM bicarbonate inside, so there is a large outward bicarbonate gradient as well as K gradient to drive the cotransporter.

The physiological function of this hitherto unknown K/HCO_3^- cotransporter is obscure. But the new technique applied by Zhao *et al.* may breathe fresh life into the study of bicarbonate transport generally. □

Roger C. Thomas is in the Department of Physiology, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, UK.

1. Zhao, J., Hogan, E. M., Bevensee, M. O. & Boron, W. F. *Nature* **374**, 636–639 (1995).
2. Maren, T. H. *Physiol. Rev.* **47**, 595–781 (1967).
3. Widdas, W. F., Baker, G. F. & Baker, P. *Cytobios* **80**, 7–24 (1994).
4. Thomas, R. C. *J. Physiol., Lond.* **273**, 317–338 (1977).
5. Schwenning, C. J. & Boron, W. F. *J. Physiol., Lond.* **475**, 59–67 (1994).
6. Boron, W. F. & Russell, J. M. *J. gen. Physiol.* **81**, 373–399 (1983).

DAEDALUS

Cold melting

MANY products and components are given their final form by casting or moulding. This fails with refractory ceramics, and also with materials that decompose on heating, such as wood, concrete, textiles, bone and leather. Daedalus is now devising a way of melting anything, simply and safely.

He points out that the molecules in a solid lattice vibrate at the Einstein frequency, which is typically around 10^{12} Hz or so. If you heat a solid, you excite this vibrational mode, which ultimately melts it. Unfortunately, you also excite all the other modes as well, including those that lead to molecular breakdown. The answer is to excite the critical lattice mode on its own. In principle, a far-infrared laser tuned to the Einstein frequency should do the trick. At first Daedalus feared that its energy would be swiftly transferred to the higher, more destructive modes. But he now thinks otherwise. Energy is rapidly shifted from higher to lower modes; but the reverse process is not so easy. Many phosphors absorb ultraviolet and re-emit visible light, but none go the other way. So the far-infrared beam should melt the target 'coldly', without thermal decomposition.

Far-infrared lasers, however, are very hard to tune. DREADCO's chemists are therefore trying a converse strategy. They are doping suitable solids, subjecting them to isotopic substitution and so on, to shift their Einstein frequencies to a given laser line. Irradiated by that line, they should sag and slump to a cool liquid puddle. Switch the laser off, and they should rapidly set, warming slightly as the absorbed energy degrades to heat.

If the idea works, more versatile infrared emitters will be pressed into service: free-electron lasers, Smith–Purcell sources and so on. These can be tuned to the Einstein frequencies of useful solids. The results will transform technology. Concrete that can be melted in place and re-set instantly; cupboards, sideboards and grand pianos cast from liquid wood; rococo moulded marble tombstones; seamless drawn leather shoes and coats, and one-piece die-cast leather footballs; paper rolled from sawdust slurry; hamburgers cast from molten meat; all will be possible.

Better still, all these products could be modified, glued, dismantled, recast or recycled by a tuned infrared beam which would cold-melt the target material and leave the rest untouched. Daedalus is now dreaming up an infrared torch to soften the wreckage around trapped crash-victims, and infrared figure reshaping to compete with plastic surgery.

David Jones

Unexpected mutagen in fish

SIR — We have previously observed powerful mutagenic activity towards *Salmonella typhimurium* TA 1535 in extracts of the Japanese fish, Sanma hiraki, treated in the laboratory with salt and nitrite at pH 3, a technique similar to the salting-pickling method for preserving fish in northern Japan, a high-risk area for stomach cancer^{1,2}. Gavage of such extracts to rats induced cancer of the glandular stomach³. It has been generally accepted that *N*-nitroso compounds are involved in the aetiology of gastric cancer⁴⁻⁶. We have now made the unexpected observation that the mutagen isolated from salt-nitrite-treated Sanma hiraki fish is not an *N*-nitroso compound, but 2-chloro-4-methylthiobutanoic acid, as verified by techniques including NMR, infrared spectrometry and low- and high-resolution mass spectrometry.

Frozen Sanma hiraki, purchased from a local Japanese food store, was homogenized (1:2, w/v) in water. After centrifugation, NaCl was added to the supernatant to give a concentration of 2%. The solution was adjusted to pH 3 with 12 M HCl, centrifuged again, and sodium nitrite was added to a concentration of 73 mM, while the pH was maintained at 3 with 12 M HCl. After 1 h at 23 °C, excess nitrite was eliminated with an equimolar amount of ammonium sulphamate. Aliquots were taken during this and subsequent procedures to monitor mutagenicity towards *S. typhimurium* TA 1535 without rat liver S9 fraction. The mutagen was extracted from the mixture with ethyl acetate. The organic phase was partitioned with 0.1 M NaOH to move the mutagen into the aqueous layer, in order to eliminate lipids. The clear, aqueous phase was adjusted to

pH 3 and extracted with methylene chloride. After removing methylene chloride *in vacuo*, the residue was dissolved in methylene chloride/hexane (1:1) and passed through a silica Maxi-Clean solid-phase extraction cartridge (Alltech, 20988). The mutagen was eluted from this cartridge with a mixture of hexane: methylene chloride: ethyl acetate

to yield the carbonium ion that in water yields the expected 2-hydroxy-4-methylthiobutanoic acid (Fig. 2). However, when chloride is present, it reacts with the carbonium ion to yield the 2-chloroderivative.

Protein hydrolysates in some commercial products contain mutagenic and genotoxic glycerol chlorohydrins^{8,9}. Salt enhances the mutagenicity of nitrite-treated extracts of black beans, perhaps through a similar mechanism, in model studies bearing on aetiological

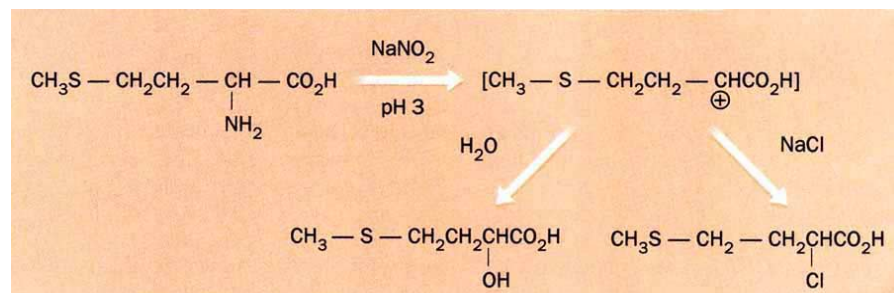


FIG. 2 Suggested scheme for the conversion of methionine to mutagen.

(60:30:10) and further purified by three steps of high-performance liquid chromatography (HPLC), giving a single homogeneous peak. As determined by gas chromatography/mass spectrometry, the mutagen had apparent molecular ion peaks at *m/z* 168 and 170, the latter with 35% intensity of the former, indicating the presence of Cl. A possible formula, C₅H₉ClO₂S, was indicated by high-resolution mass spectrometry. The NMR spectrum showed signals consistent with the presence of one CH₃, two CH₂ and one CH group. The assignments were confirmed by decoupling experiments. There were no aromatic hydrogens. The infrared spectrum displayed a carbonyl group (1,726 cm⁻¹). Evaluation of the combined results revealed the fish mutagen to be 2-chloro-4-methylthiobutanoic acid (CMBA), which has not been previously described in *Chemical Abstracts*. The chemical is a slightly viscous, colourless liquid.

The structure suggested that the mutagen might be derived from methionine. Therefore, L-methionine was taken through the same reactions. The HPLC procedures, as applied to the fish extract, gave a peak with a retention time identical to that of the mutagen from the fish. Mass spectral and NMR data confirmed the identity of the product as CMBA, providing unambiguous, independent support for the structure of the mutagen isolated from fish (see Fig. 1). La Vecchia *et al.*⁷ found a relative risk of gastric cancer of about 2.40 for the highest quintile of dietary methionine intake in a northern Italian population.

Omission of sodium chloride during the reaction failed to produce the mutagen. The reaction of the amino group of methionine with nitrite generates a reac-

tive carbonium ion that in water yields the expected 2-hydroxy-4-methylthiobutanoic acid (Fig. 2). However, when chloride is present, it reacts with the carbonium ion to yield the 2-chloroderivative.

Protein hydrolysates in some commercial products contain mutagenic and genotoxic glycerol chlorohydrins^{8,9}. Salt enhances the mutagenicity of nitrite-treated extracts of black beans, perhaps through a similar mechanism, in model studies bearing on aetiological

factors for gastric cancer in Latin America⁹. Monochloroacetic acid, the simplest compound in this series, is neither mutagenic¹⁰ nor carcinogenic in mice and rats¹¹. Thus, the appreciable mutagenicity and activity in the DNA repair test in hepatocytes of Williams of 2-chloro-4-methylthiobutanoic acid indicate that this molecule has genotoxic attributes based on the possible formation of reactive carbonium or sulphonium ions.

Thus, treatment of fish with both salt and nitrite generates a novel mutagen, an α-chlorobutanoic acid. We are now studying the role of this compound in the induction of gastric cancer.

Wei Chen

John H. Weisburger*

Emerich S. Flala

Steven G. Carmella

Di Chen

Thomas E. Spratt

Stephen S. Hecht

*American Health Foundation,
Valhalla, New York 10595, USA*

*To whom correspondence should be addressed.

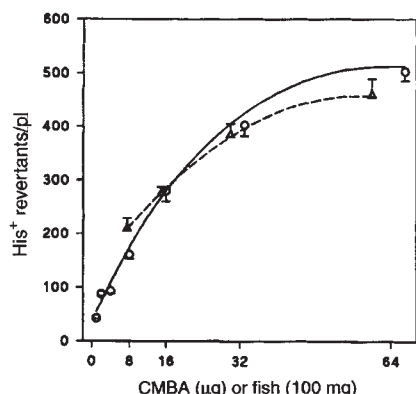


FIG. 1 Direct-acting mutagenicity in *S. typhimurium* TA 1535 produced by 2-chloro-4-methylthiobutanoic acid (µg CMBA, three plates at each dose level; circles) or extracts of Sanma hiraki fish treated with salt and nitrite at pH 3 (see text; data shown from four plates at each dose level in units of 100 mg wet weight fish; triangles). Error bars are $\pm$ s.d.

- Marquardt, H.F., Rufino, F. & Weisburger, J.H. *Science* **196**, 1000-1001 (1977).
- Binkerd, E.F. & Kolari, O.E. *Fd Cosmet. Tox.* **13**, 655-661 (1975).
- Weisburger, J.H., Marquardt, H., Hirota, N., Mori, H. & Williams, G.M. *J. natn. Cancer Inst.* **64**, 163-167 (1980).
- Wakabayashi, K., Nagao, M. & Sugimura, T. *Cancer Surv.* **8**, 385-399 (1989).
- O'Neill, I.K., Chen, J. & Bartsch, H. (eds) *IARC Sci. Pub.* **105** (Int. Ag. Res. Cancer, Lyon, 1991).
- Forman, D. & Shuker, D.E.G. (eds) *Cancer Surv.* **8**, 201-492 (1989).
- La Vecchia, C., Ferraroni, M., D'Avanzo, B., Decarli, A. & Franceschi, S. *Cancer Epid. Biomarkers Prev.* **3**, 393-398 (1994).
- Plasecki, A., Ruge, A. & Marquardt, H. *Arzneimittel-Forsch.* **40** (II), 9, 1054-1055 (1990).
- Rojas-Campos, N., Sigaran, M.F., Bravo, A.V., Jimenez-Ulate, F. & Correa, P. *Nutr. Cancer* **14**, 1-3 (1990).
- Ashby, J. & Tennant, R.W. *Mutagenesis* **9**, 7-15 (1994).
- National Toxicology Program Tech. Rep. 896 (NIH Publ. 92-2851, 1992).

Exohedral adducts of La@C₈₂

SIR — We recently reported a means of derivatizing fullerenes C₆₀ and C₇₀ by photochemical attachment of disilirane (disilacyclopropane)^{1,2}. Here we show that this approach can be used to prepare an exohedral derivative of the endohedral metallofullerene La@C₈₂.

We synthesized and purified La@C₈₂ by a method we developed recently³. A toluene solution of La@C₈₂ and 1, 1, 2, 2-tetrakis(2, 4, 6-trimethylphenyl)-1, 2-disilirane ((Mes₂Si)₂CH₂; Mes = 2, 4, 6-trimethylphenyl) (**1a**) was photoirradiated at 20 °C with a tungsten-halogen lamp (cutoff <400 nm) in a degassed sealed tube. The reaction product was analysed by laser-desorption time-of-flight spectrometry (LD-TOF), fast-atom bombardment (FAB) mass spectrometry, electron paramagnetic resonance (EPR) and visible/near-infrared spectroscopy. The Fourier-transform infrared (FTIR) spectrum and the time-dependent changes in the EPR spectra are available as Supplementary Information*.

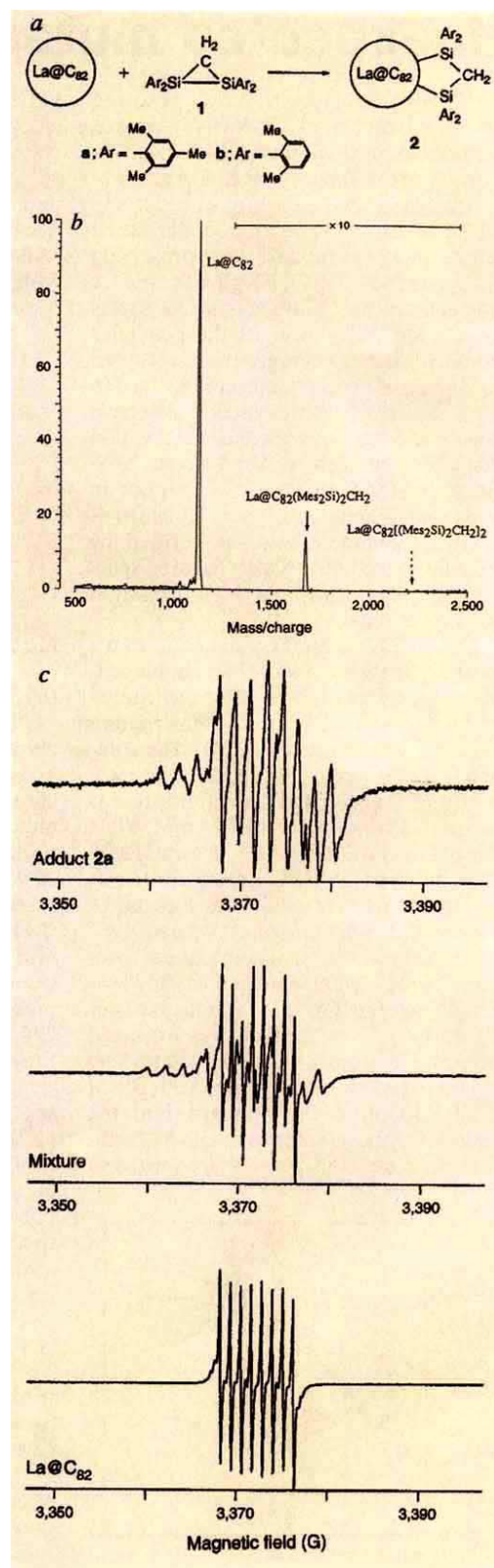
The LD-TOF mass spectrum of the product shows the presence of La@C₈₂(Mes₂Si)₂CH₂ (**2a**) (*b* in the figure). No molecular ion peaks ascribable to multiple-addition products such as La@C₈₂[(Mes₂Si)₂CH₂]₂ were observed. The positive-ion FAB mass spectrum also confirms the formation of **2a**. The observed ion intensity ratio of a group of peaks at mass/charge ratios of 1,669–1,672 for **2a** agrees with the carbon and silicon isotope distributions⁴. In both mass spectra, the ion peak for La@C₈₂ serves as a reference. The LD-TOF and FAB mass analyses of the exohedral derivatives of empty fullerenes, C₆₀(Mes₂Si)₂CH₂ (ref. 1) and C₇₀(Mes₂Si)₂CH₂ (ref. 2), usually show strong fragmentation peaks such as C₆₀ and C₇₀, respectively, due to a loss of exohedral functional group [(Mes₂Si)₂CH₂] from the adducts. Very similar results were also obtained with 1, 1, 2, 2-tetrakis(2, 6-dimethylphenyl)-1,

2-disilirane (**1b**) (*a* in the figure).

The EPR spectrum of the adduct (**2a**) in 1, 2, 4-trichlorobenzene is compared in *c* in the figure with that of La@C₈₂ (refs 3, 5) which contains one set of equally spaced octet lines with a hyperfine splitting of 1.15 G. Several new sets of octet lines are observed in the spectrum of the adduct. We also measured the EPR spectrum of a mixture of both species during the reaction, in which octet lines of both La@C₈₂ and **2a** are seen. Eight equally spaced, broad lines with a hyperfine coupling (HFC) constant of ~1.7 G are followed by another broad set of octet lines (HFC constant ~1.8 G) of weaker intensity. These spectra show the formation of at least two product species with a different La isotropic splitting, presumably two positionally isomeric forms of the disilirane derivative. The electronic structure of La@C₈₂ can be formally written as La³⁺C₈₂³⁻, three valence electrons on La being transferred to the LUMO and LUMO+1 of C₈₂ (ref. 6). The relatively large HFC constants of **2a** compared with that of La@C₈₂ suggest that the La atom is less positively charged in **2a** than in La@C₈₂. It has been shown⁷ that silicon-containing substituents donate almost one electron to fullerene cages, raising the LUMO and LUMO+1 levels by about 0.5 eV. Thus, one would expect the degree of electron transfer from La to the silicon-derivatized fullerene cage to be significantly suppressed.

The visible and near-infrared absorption spectrum of **2a** in 1, 2, 4-trichlorobenzene (available as Supplementary Information) does not contain the absorption peaks at 640, 1,010 and 1,430 nm observed for La@C₈₂ (refs 3, 5). However, both La@C₈₂ and **2a** show broad near-infrared absorption bands down to 2,000 nm because of their open-shell electronic structures.

We found that La@C₈₂ is also thermally reactive towards **1**. When a toluene solution of La@C₈₂ and **1** was heated at 80 °C for 10 h, formation of the adduct **2a** was verified by means of FAB mass spectrometry and EPR. The high thermal reactivity of La@C₈₂ towards **1** can be rationalized on the basis of its stronger electron donor and acceptor properties⁸ relative to C₆₀ and C₇₀: *ab initio* calculations⁸ predict that the ionization potential (6.19 eV) and electron affinity (3.22 eV) of La@C₈₂ are much smaller and larger than those for C₆₀ (7.78 and



a, Reaction scheme for preparation of the exohedral adducts **2a** and **2b**. *b*, Laser-desorption time-of-flight (LD-TOF) mass spectrum of the adduct La@C₈₂(Mes₂Si)₂CH₂ (**2a**). Laser desorption and ionization were done at 337 nm. *c*, EPR spectrum of the adduct La@C₈₂(Mes₂Si)₂CH₂ (**2a**) (top), the mixture of La@C₈₂ and **2a** during the reaction (middle), and La@C₈₂ (bottom). The samples in 1,2,4-trichlorobenzene were degassed several times by a freeze-thaw cycle and sealed in a thin-wall quartz tube. EPR measurements were performed at ambient temperature.

1. Akasaka, T., Ando, W., Kobayashi, K. & Nagase, S. *J. Am. chem. Soc.* **115**, 10366–10367 (1993).
2. Akasaka, T., Mitsuhashi, E., Ando, W., Kobayashi, K. & Nagase, S. *J. Am. chem. Soc.* **116**, 2627–2628 (1994).
3. Yamamoto, K., Funasaka, H., Takahashi, T. & Akasaka, T. *J. phys. Chem.* **98**, 2008–2011 (1994).
4. Shinohara, H. *et al. Rapid Commun. Mass Spectrom.* **6**, 413–416 (1992).
5. Kikuchi, K. *et al. Chem. Phys. Lett.* **216**, 67–71 (1993).
6. Nagase, S. & Kobayashi, K. *Chem. Phys. Lett.* **214**, 57–63 (1993).
7. Suzuki, T. *et al. J. Am. chem. Soc.* **116**, 1359–1363 (1994).
8. Nagase, S. & Kobayashi, K. *J. chem. Soc. chem. Commun.* 1837–1838 (1994).
9. Suzuki, T. *et al. J. Am. chem. Soc.* **115**, 11006–11007 (1993).
10. Saunders, M. *et al. J. Am. chem. Soc.* **116**, 3621–3622 (1994).

*Requests for Supplementary Information should be addressed to Mary Sheehan at the London editorial office of *Nature*.

2.57 eV) and C₇₀ (7.64 and 2.69 eV), respectively. This is consistent with the fact that La@C₈₂ has low oxidation and reduction potentials relative to those of C₆₀ and C₇₀ (ref. 9). Even though the electron affinity of C₈₂ (3.37 eV) is comparable to that of La@C₈₂, thermal addition of disilirane to C₈₂ is suppressed owing to its larger ionization potential (6.96 eV)⁸.

An exohedral derivative of an endohedral fullerene compound was reported previously by Saunders and co-workers¹⁰, who allowed helium atoms to perforce through the cage of a derivative of empty C₆₀. In contrast, the compounds reported here are made by attaching substituents to a preformed endohedral species. We anticipate that this approach may enable us to 'tune' chemically the degree of charge transfer in, and thus the properties of, the metallofullerenes.

Takeshi Akasaka

Department of Chemistry,
University of Tsukuba,
Tsukuba, Ibaraki 305, Japan

Tatsuhisa Kato

Institute for Molecular Science,
Myodaiji, Okazaki 444, Japan
Kaoru Kobayashi, Shigeru Nagase
Department of Chemistry,
Faculty of Education,
Yokohama National University,
Yokohama 240, Japan

Kazunori Yamamoto, Hideyuki Funasaka Takeshi Takahashi

Nuclear Fuel Technology Development
Division, Tokai Works,
Power Reactor & Nuclear Fuel Development
Corporation,
Tokai, Ibaraki 319-11, Japan

No UVB effect?

SIR — McMinn *et al.*¹ have examined the relative abundances of diatom taxa in three sediment cores, each taken from a different fjord in the Vestfold Hills region of eastern Antarctica. Over a 20-year period they noted little change in species composition based on relative abundances, and they thus concluded "that the enhanced ultraviolet-B levels resulting from the 'ozone hole' [have] had little effect on the diatom component of the phytoplankton community."

McMinn *et al.* acknowledge that their results probably apply only to coastal diatom communities largely shielded from ultraviolet-B radiation by thick ice cover. Unfortunately, the title of their letter, "Minimal effects of ultraviolet-B radiation on Antarctic diatoms over the past 20 years", implies something much broader for the Antarctic ecosystem. The data reported in their paper do not warrant this claim relative to Antarctic diatoms in general, or perhaps even for the diatom communities from the Vestfold Hills region.

First, the diatom communities they have studied represent only a fraction of the diatom-based productivity of the Southern Ocean. It is the planktonic ice-edge and sea-ice communities of the more northern areas of the Southern Ocean that are the most productive during spring, are exposed to higher levels of ultraviolet-B, and are therefore more vulnerable to ozone depletion events. Second, the results are less compelling because the diatom counts were done on a relative rather than an absolute basis. For example, there could have been a 90% decrease (or increase) in the absolute abundance of diatom cells in Antarctic waters, without a change in the relative abundance of species. Further, large interannual variation in absolute diatom abundance could mask subtle but significant shifts in the relative proportions of the community if the counts were not statistically stratified.

Third, the potential effects of ultraviolet-B on diatom community species richness and diversity are problematic. The notion that environmental stresses such as ultraviolet-B will reduce species numbers and diversity is not necessarily correct. A long-term *in situ* study² showing that diatom community diversity increased under exposure to ultraviolet-B demonstrates the weakness of diversity as a measure of community response to this parameter. We believe that the absence of a decline in diatom species richness or diversity over the past 20 years in the data collected by McMinn *et al.* does not necessarily indicate that "increased UVB irradiances have had little effect on the planktonic diatom community."

Fourth, Smith *et al.*³ showed that declines in primary productivity under the influence of elevated ultraviolet-B were less than the interannual variability of productivity. Unless there were distinct taxonomic patterns reflecting interannual variability evident in the McMinn *et al.* core data before the ozone depletion events initiated in the 1970s, there is no reason to suspect that core data would provide any insight into ultraviolet-B effects.

Last, the statement that levels of predation (grazing) are unlikely to have changed significantly over the past 20 years is not supported by any data (as acknowledged by McMinn *et al.*). Recent research in shallow, freshwater habitats⁴ has indicated that grazing populations can be more susceptible to ultraviolet-B damage than primary producers and altered trophic-level interactions can produce greater effects on diatom communities than can direct effects of ultraviolet-B. It is now clear that changes in diatom numbers or species alone do not adequately address the issue of ecosystem change. The impact of ozone depletion on the Antarctic ecosystem is

still very uncertain. Our concern is that the limited data provided by McMinn *et al.* do not substantiate the implication of the title of their paper.

M. L. Bothwell

Environmental Sciences Division,
National Hydrology Research Institute,
Saskatoon, Saskatchewan S7N 3H5,
Canada

D. Karentz

Department of Biology,
University of San Francisco,
San Francisco,
California 94117-1080, USA

E. J. Carpenter

Marine Sciences Research Center,
State University of New York,
Stony Brook,
New York 11794-5000, USA

1. McMinn, A., Heljas, H. & Hodgson, D. *Nature* **370**, 547–549 (1994).
2. Bothwell, M. L., Sherbot, D., Roberge, A. C. & Daley, R. *J. Phycol.* **29**, 24–35 (1993).
3. Smith, R. C., *et al.* *Science* **255**, 952–959 (1992).
4. Bothwell, M. L., Sherbot, D. M. J. & Pollock, C. M. *Science* **265**, 97–100 (1994).

Diabetes genes — *mutatis mutandis*

SIR — In your recent editorial ("Nature Sister Publications", *Nature* **374**, 95; 1995), there were several errors in the description of our paper reporting a mutation predisposing to type 1, or insulin-dependent diabetes mellitus (IDDM).

Insulin-dependent diabetes mellitus is never abbreviated as "IDS", for example, nor is the HLA complex on chromosome 6 ever referred to as the "LA complex". The susceptibility locus on chromosome 11p15 is called IDDM2, and this locus, which we (*Nature Genetics* **9**, 284–292; 1995) identified as the variable number of tandem repeats locus within the transcriptional regulatory 5' region of the insulin gene is never referred to as a "VENTURE". It is commonly referred to as a VNTR, and this correct abbreviation is used later in the editorial.

In our paper, we used families from the United Kingdom, United States and Denmark, but not from France. One of our interesting results was the finding that the 698-base-pair allele of the VNTR appears to encode different susceptibility from other VNTR alleles. This allele has 34 repeats, not 50 as claimed in the editorial. This is of key importance, as the number and DNA sequence of the repeats appears to be a critical factor for disease susceptibility and insulin gene expression (see also Kennedy *et al.*, *Nature Genetics* **9**, 293–298; 1995).

In the same issue of *Nature Genetics*, Hager *et al.* (**9**, 299–304; 1995) report that a mutation of the glucagon receptor gene is associated with type 2, or non-insulin-dependent, diabetes mellitus in

France. It was stated in the same editorial that these authors have also obtained statistically significant results from Sardinia, but this is not true. In effect, the result remains to be replicated in independent case control and family studies, as is the case for most genetic associations in type-2 diabetes published so far.

John A. Todd, Simon T. Bennett

The Wellcome Trust Centre for Human Genetics,

University of Oxford, Oxford OX3 7BN, UK

Controlled-size nanocapsules

SIR — Encapsulation of second phases inside graphite shells/nanotubes^{1–7} is of considerable significance not only because it offers an opportunity to investigate dimensionally confined systems, but also because the encapsulated species are likely to be immune to environmental effects or degradation owing to the protective graphite sheets around them⁸. It has been difficult to encapsulate ferromagnetic transition metals using the traditional graphite arc process, however, because of poor and irreproducible yield, no particle size control and the unavoidable presence

of unwanted carbides, graphite flakes, nanotubes and particulates.

Here we report a novel, large-scale synthesis method to encapsulate controlled-size ferromagnetic transition metals, such as Fe, Co and Ni, inside protective graphite shells. The experimental set-up consists of an arc chamber with a vertical tungsten cathode (about 3.2 mm diameter), and an anode composed of a metal block placed inside a graphite crucible with an inner diameter of about 38 mm. A transverse helium gas jet passes through the arc plasma and provides quenching of the metal vapour. The metal block (such as nickel) is usually thermally isolated from the graphite crucible by a thin *c*-axis-oriented graphite foil. Under arc evaporation conditions, the metal block forms a molten pool and the spatial distribution of the arc plasma is wide enough to engulf part of the surrounding graphite crucible and foil, both of which provide just enough supply of carbon for subsequent graphitic encapsulation.

A well-defined log-normal size distribution with mean particle size ranging from 7 to 14 nm has been obtained by controlling the helium jet velocity between 56 and 20 m s⁻¹ (ref. 9). The resultant product consists mainly of nanocrystalline metal particles, graphite-encapsulated nanocrystalline metal with a minority of amorphous carbonaceous debris. It is important to note that the synthesis product in our method is virtually devoid of any graphitic flakes, nanotubes or coarse nanoparticles, which commonly occur in the traditional graphite-graphite arc process and are very difficult to eliminate.

The condensed product is collected and washed (under exposure to ultrasound) in a strong solvent (such as HCl or aqua regia: a 3:1 mixture of HCl and HNO₃), which dissolves unprotected metallic particles, leaving behind a large amount (about 20–40%, but sometimes as large as 70%) of residue containing graphite-encapsulated nanocrystals. The acid-washed residue consists of graphite-encapsulated nanocrystals and amorphous carbonaceous debris, with no trace of graphite nanotubes or graphite nanoparticles. The minor amount of carbonaceous debris present may be oxidized and removed using a method similar to that reported by Ebbesen *et al.*¹⁰ for the purification of carbon nanotubes. The final yield depends on the type of metal species and critically on some operating variables such as the geometry of the graphite crucible-metal anode assembly (for example, inner diameter of the graphite crucible).

Figure 1 shows a dispersed ensemble of nanocrystalline nickel particles completely encapsulated with graphite. X-ray microanalysis, electron nanodiffraction and electron energy loss spectroscopy fine-structure analysis confirmed that the encapsulated product is metallic nickel and that the coating is graphite. The intimate and contiguous graphite fringes around the nanocrystal particles is evidence for complete encapsulation by graphite shells.

Measurements of the magnetic properties of the encapsulated nanocrystalline nickel indicate that the intrinsic properties of ferromagnetic nickel are retained even with encapsulation in graphite shells. Saturation magnetization, Curie temperature, and magnetization versus field loops are consistent with those of nanocrystalline nickel (which also shows superparamagnetic behaviour below about 4–6 nm particle size). Figure 2 shows a plot of magnetization against temperature for graphite-encapsulated nickel. The Curie temperature is 358 °C, with an error of about 10 °C.

Our results demonstrate successful encapsulation of controlled-size ferromagnetic nanocrystalline metals which retain their intrinsic nanocrystal properties. Graphite encapsulation makes the nanocrystalline metal particles immune to environmental degradation: they are insoluble and undamaged even when exposed to aqua regia. Some encapsulated nanocrystals were kept in aqua regia for months, but did not dissolve, and the protective graphite coating remained intact.

The approach presented here not only results in a much higher yield of controlled-size encapsulated species, but also the occurrence of carbides and carbonaceous debris (for example, graphite flakes, buckytubes and their particulate derivatives), which is very difficult to eliminate in subsequent processes, is minimized. The synthesis approach reported here may be extended to other materials that catalyse nucleation of the graphite form of carbon on their surfaces.

Vinayak P. David, Jonathon J. Host

M. H. Teng, Brian Elliott

Jinha Hwang, D. Lynn Johnson

Thomas O. Mason, Julia R. Weertman

Department of Materials Science & Engineering and Materials Research Center,

Northwestern University,

Evanston, Illinois 60208, USA

1. Ajayan, P.M. & Iijima, S. *Nature* **361**, 333–334 (1993).

2. Ajayan, P.M. *et al.* *Nature* **362**, 522–525 (1993).

3. Ruoff, R. *et al.* *Science* **259**, 346–348 (1993).

4. Seraphin, S. *et al.* *Nature* **362**, 503–503 (1993).

5. Tomita, M., Saito, Y. & Hayashi, T. *J. appl. Phys.* **32**, L280–L282 (1993).

6. Majetich, S.A. *et al.* *Phys. Rev. B* **48**, 16845 (1993).

7. Saito, Y. *et al.* *J. Phys. Chem. Solids* **54**, 1849 (1993).

8. Saito, Y. in *Fullerenes: Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials* (eds Kadish, K. M. & Ruoff, R. S.) 1419–1432 (Electrochem. Soc., New Jersey, 1994).

9. Teng, M. H. *et al.* *J. Mater. Res.* **10**, 233–236 (1995).

10. Ebbesen, T.W., Ajayan, P.M., Hiura, H. & Tanigaki, K. *Nature* **367**, 519–519 (1994).

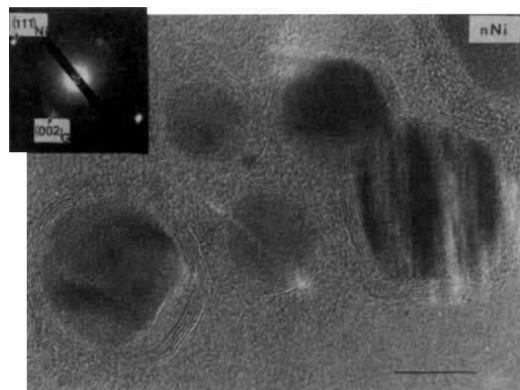


FIG. 1 High-resolution electron microscope image of graphite-encapsulated nanocrystalline nickel particles. Inset, nanodiffraction pattern indicating graphite (002) and nickel (111) reflections. Scale bar, 10 nm.

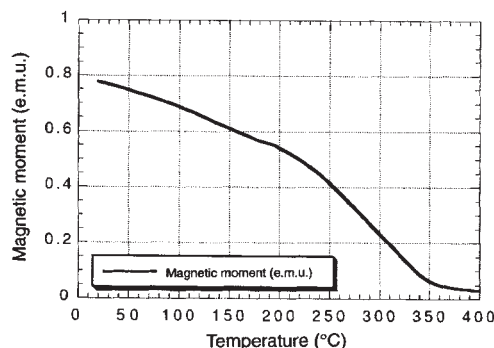


FIG. 2 Magnetization versus temperature plot of graphite-encapsulated nanocrystalline nickel. The Curie temperature determined from the plot (~358 °C) is consistent with that of metallic nickel.

Shock of the new

David Edgerton

American Technological Sublime. By David E. Nye. MIT Press: 1994. \$35, £31.50.

FROM the late nineteenth century to today, the United States has been the most efficient large industrial economy in the world. At the peak of its relative power Americans were some 30 years ahead of Europeans in the ownership of consumer durables. It has been, without question, the great technological power of the twentieth century. Industrially and technologically it has stood for modernity, even though it subverts our common cultural and political picture of modernity. Its religious observance is astonishingly high; it has never had a mass socialist party; its progressivist ideologies have differed markedly from European models. Many have found in Americans' love affair with modern technology a distinct 'American ideology'.

Some have pointed this out in order to shame their own country into a greater technological enthusiasm; others have sought to criticize Americans' infatuation with things. Critics have long bemoaned the grip that 'technological determinism' has had on the US imagination. They argue this simply hid from view the actions of powerful human forces such as corporations. But although many have criticized Americans' over-emphasis on the independent role of technology in shaping society and history, few have examined their actual attitudes to technology and how they have changed over time.

In his fine book David Nye explores one fragment of that story: America's celebration of large-scale and powerful technologies. He looks in particular at things he labels 'sublime': those that inspired awe, terror and astonishment even on repeated viewings. Nye makes very clear that Americans did experience some technologies in this way, indeed in ways similar to some of the great natural wonders such as the Grand Canyon. The canal that linked New York with the Great Lakes, the first railroads, great bridges, skyscrapers, the atomic bomb, space travel, were all surprisingly public spectacles that gripped the American mind. Bridges such as Brooklyn and the Golden Gate attracted huge attention (and indeed continue to do so). After 1945 Americans made long trips to see atomic and hydro-

gen bombs being exploded, and indeed complained that the explosions did not get ever-more spectacular.

Nye shows we understand the same thing in different ways over time, and different people understand things in different ways. The thing in itself did not determine Americans' reaction to it; interpretations of something as fixed as the Grand Canyon have changed,



Topping it all: a worker constructing the Chrysler Building in Manhattan in 1928. For most of this century the island has boasted the world's tallest habitable building.

betokening changes in attitudes as a whole. A seventeenth-century European might have seen it as nature at its most barbarous, whereas a nineteenth-century American would have seen in its awesome sublimity the work of God. Today, Nye tells us, the most common questions asked of the park wardens assume that the canyon was made by human beings.

The Grand Canyon, even if thought of as the result of modern technology, continues to attract visitors, but there have been great changes in what counts as technologically sublime. Novelty became a key feature. Only the new came to be regarded as sublime: the eyes of a whole

city would turn skywards when the first aeroplanes appeared, but only the equivalent of train-spotters would do the same today. Bridges, skyscrapers and atomic explosions had to be bigger than the last to be worth seeing; America came to be in love with technological change rather than with technology.

To judge from international exhibitions, the kind of thing presented and what was deemed sublime also changed. In the nineteenth century the products of industry of nations were the attraction; by the early twentieth century it was the process of production (notably the assembly line) that filled people with awe. Indeed, Ford and many other companies welcomed visitors, in very large numbers, to their plants. At the New York World's Fair (1939-40) current products and processes were rather *passé*: it was the future that was yet to be created that drew the crowds and filled the stands put up by massive US corporations. The technological sublime was the as yet unrealized creation of the corporate research laboratory. It is not accidental, therefore, that British television's longest running programme on technology should be called "Tomorrow's World".

Nye points to another change: in the nineteenth and early twentieth centuries, the celebration of big technologies was a celebration of the republic itself; it was a civic act. By the 1930s it had become corporate propaganda.

Nye brings us up to date with a view of Las Vegas, an indiscriminate rolling together of the natural and technological sublimities of the past. The new Las Vegas incorporates not only the electrical sublime of the early twentieth century, but also ersatz volcanoes, waterfalls and skyscrapers. It is not nature's wonders, or man's production, or indeed the future, that is celebrated in Las Vegas but a dream world of

consumption.

Nye's book focuses on what he recognizes is the untypical case of the United States, and on only one aspect of Americans' attitudes to technology. But the history of popular thinking about technology, as Nye's book shows, can help us to understand the limits of our own understanding of the relationship between technology and society, and indeed how those limits have arisen. □

David Edgerton is at the Centre for the History of Science, Technology and Medicine, Imperial College, Sherfield Building, London SW7 2AZ, UK.

Botanical roots

Sandra Knapp

A Passion for Plants: From the Rainforests of Brazil to Kew Gardens. By Clive Langmead. *Lion*: 1995. Pp. 201. £16.99.

THERE are basically three types of biographies: those about politicians and other public figures written during their careers; those, often about scientific figures, written after a person's death and reviewing his or



From A Passion For Plants

Man of vision — Ghilleen Prance doing field work.

her life and achievements; and autobiographies, which provide a rare glimpse into the working of someone else's mind. This book is of the first sort. It is about the life and vision of Ghilleen Prance the director of the Royal Botanical Gardens at Kew, rather than about Ghilleen Prance the plant taxonomist. I am sure that this was the intention of the author, but I am left with the feeling that from a man as driven and as interesting as Prance there is still more to come.

The life described here is a fascinating one, full of adventure and hard work. The detail with which Prance's early life is portrayed makes the reader see the director of Kew as a man rather than a figurehead, although anyone who has had any contact with Prance already knows that he is a remarkably accessible, open and friendly figurehead. The emphasis on the pervasive influence of his strong Christian beliefs on his life both in and out of science (lives that in his case are not all that separable) is due to the importance with which he regards these beliefs and to the fact that the publisher specializes in Christian matters. At one point in his life, Prance considered ordination but was persuaded by his future

father-in-law to pursue a career in botany and not to waste the gift he had for natural history.

Taxonomy and systematics are often seen as dull, parochial and irrelevant in this age of molecular biology and gene manipulation. With science very much being pushed towards wealth creation, we are in danger of losing sight of the true curiosity that drives the quest for scientific knowledge. This book, by virtue of being for a nonspecialist audience, shows clearly that the frontiers of discovery are not all in laboratories — they can also be found in the Amazon. The story of the pollination of

Victoria amazonica, the giant Amazonian water lily, is a classic of botanical detective work. (The sequence was beautifully portrayed in David Attenborough's recent television series "The Private Life of Plants".) To discover that the beetle pollinators are trapped inside the flower for 18 hours while pollen is released, then let free to cross-pollinate other younger flowers, is neither trivial nor irrelevant; knowledge of such complex interactions and their mechanics and distribution will ultimately allow us to understand the world's great ecosystems. Sustainable development, the buzzword of the moment, can be achieved only by knowing what is in an area and how it works.

Knowledge about the names, relationships and interactions of plants and animals is the basis for all biology and is of crucial importance to the future of our own and other species. Prance has spent his career working from within to further these views and to make the plight of rainforests, particularly those of the Amazon, understandable to the 'man in the street'. His decision not to co-author with Robert Goodland *Green Hell to Red Desert*, one of the first books to warn of the threat of Amazonian rainforest development, allowed him to continue to work in Brazil, and to bring about changes he felt were essential to the future of the region.

As director of Kew, Prance has been able to spread his botanical message to an even broader audience. His decision to leave the New York Botanical Garden, where he had worked for 24 years, and come to Kew is portrayed as a decision for the good of botany rather than a personal career move. The directorship of Kew may well be the top botany job in the world, but his use of the position as a platform for educating people in all walks of life about the current grave environmental crisis shows great courage and dedication. I first met Prance while he was at New York, and his influence on my generation of botanists went far beyond the boundaries of the New York Botanical Garden.

Clive Langmead has produced an easy to read, interesting and somewhat eulogistic biography of an exceptional person. But while reading the book I found myself constantly annoyed by the lack of a personality shining through. I hope that one day Prance will write an autobiography and give us all the chance to see his mind at work. For I suspect that, like Richard Spruce, the famous Amazonian explorer to whom he is often compared, his own thoughts about his life are infinitely more interesting than someone else's. Prance is still young as scientists go, and his future years are sure to be just as surprising and fruitful as those that have passed. □

Sandra Knapp is in the Department of Botany, The Natural History Museum, Cromwell Road, London SW7 5BD, UK.

Guide to Flowering Plant Families. By Wendy B. Zomlefer. *University of North Carolina Press*: 1994. Pp. 430. \$55 (hbk), \$27.50 (pbk).

Good illustrations and clear and concise descriptions of the families of flowering plants are hard to find in one place. Time and again, I have gone back to my bible for this sort of synoptic information, Hutchison's *Families of Flowering Plants*, published in 1934 (with a second edition in 1959).

Zomlefer's book very neatly fills a gaping hole in the literature for the student of plant taxonomy and general botany. The book has a lucid and well written introduction to modern taxonomic methods, complete with a nonthreatening treatment of cladistics and its use in taxonomy. The families of flowering plants in the United States (including Florida — that is, the subtropics) are then described, discussed and beautifully illustrated. The drawings alone make this book a treasure; they are clear, large and display all the diagnostic features necessary for the identification of plants in any particular family. The references included with each family description lead the student to the primary literature, and recent reorganizations and major taxonomic changes are fully discussed, such as those for the mints and verbenaceae (see P. Cantino, *Ann. Missouri Bot. Gard.* **79**, 361; 1992).

The book is designed for the North American market, but would be useful anywhere. The illustrated glossary will make it invaluable for anyone trying to come to grips with the sometimes obscure world of botanical terminology (drawing of a trigonous stem jumps off the page). It is rare that a textbook is also an important reference work for the professional botanist, but Zomlefer has made the information presented here so accessible, authoritative and up to date that anyone enthusiastic about plants should have the volume on their shelves.

S. K.

Stellar abundance

Stuart Ross Taylor

An Introduction to Cosmochemistry. By Charles R. Cowley. Cambridge University Press: 1995. Pp. 490. £50, \$79.95 (hbk); £19.95, \$29.95 (pbk).

COSMOCHEMISTRY is a science that has flowered only in the past 50 years as the wealth of information from geochemistry, planetology, astronomy and astrophysics has enabled us to contemplate the chemistry of the cosmos. We are now able to understand the composition and chemical evolution of much of the observable Universe. In an astonishing synthesis of nuclear physics and astronomy, the origin of the chemical elements themselves, surely one of the great intellectual triumphs, was explained a generation ago. The composition of interstellar matter and stars has slowly become intelligible. Buried within meteorites that we can examine in terrestrial laboratories are isotopic signatures of events that happened before the formation of the Sun. By now we understand much about the origin and evolution of the Solar System, from the early separation of the solar nebula from a molecular cloud to the formation of the Moon with its strange and unique composition. I therefore opened this book with a great deal of anticipation. Regrettably, it did not live up to the promise of its title.

The book is really a mixture of geology and astronomy with only a thin bridging section on cosmochemistry. The author, who in 1970 wrote *The Theory of Stellar Spectra* (Gordon and Breach), is an expert both on that topic and on the composition of peculiar stars. Astronomers will find the sections on the interpretation of stellar spectra and stellar abundances, clearly the area of most interest to the author, a useful summary, but the information is too detailed for most geochemists. In turn, the geologists and geochemists among the advanced undergraduates and graduate students at whom the book is aimed will find the geology too simple and the astronomy too difficult. Most of the geochemical data come from old or secondary sources. The early sections contain brief summaries of mineralogy, petrology, thermodynamics and isotope geology. It is not really clear what purpose these sections serve except perhaps to introduce the subject to astronomers. Better treatments are available in standard texts, and these sections could well be deleted in subsequent editions. What is missing is an in-depth discussion of geochemistry and cosmochemistry as currently practised. Only three rather brief chapters cover these topics. These sections give only a short summary of a rapidly developing field and the author sometimes loses his

way. He repeats the familiar error that the material in the Moon came from the mantle of the Earth, although current work derives it from the mantle of the impactor. This model, motivated by the failure of the lunar compositional data to match that of the terrestrial mantle, was spectacularly confirmed by the Clementine Mission.

When the author starts to deal in later chapters with his own speciality, stellar spectra, he leaps into very detailed discussions that will rapidly lose most students. These sections are written at a considerably more advanced level than those on geology, petrology and mineralogy. But in many cases, just as the discussion is becoming interesting, the reader is abandoned to consult the original literature. Curiously there is no mention that I could find of the famous discrepancy between the meteoritic and solar iron abundances that bedevilled geochemists ever since Urey and that has only recently been resolved. Holweger and colleagues at Kiel, using the Fe II lines in the solar spectra, established the equivalence of the iron abundances in the Sun and the type 1 carbonaceous chondrites. That was a considerable relief to cosmochemists, as cobalt and nickel, the two elements most closely associated with iron, had never shown any discrepancy.

The text often comes across as a pedagogical exercise, and this is heavily reflected in the style. The author seems most at

ease with sentences beginning with such well worked phrases as "let us consider a well mixed gas" and then launches into a detailed mathematical treatment. The elegance of this approach, although well suited to a class exercise, is unfortunately diminished by the many uncertainties surrounding the subject.

The dangers of moving outside one's sphere of expertise are sometimes apparent. The author is reduced to accepting, at face value, many controversial topics in geochemistry, on occasion being moved to lament that "it is beyond the competence of the present writer to comment on the relative merits of the different approaches", for example, on the composition of the Earth. Indeed, the complexity of that topic might daunt the most accomplished geochemist. But writers of textbooks, having surveyed the relevant literature, are in a better position than the average reader of their works to form an opinion. Without it, textbooks become merely a trackless listing of current problems that readers must attempt to unravel without guidance. Of course, the solution is that the experts in the particular disciplines should write the books, just as Jeans, Eddington and Hubble did earlier this century for astronomy. □

Stuart Ross Taylor is in the Department of Nuclear Physics, Research School of Physical Sciences, Australian National University, Canberra, ACT, Australia 2600.

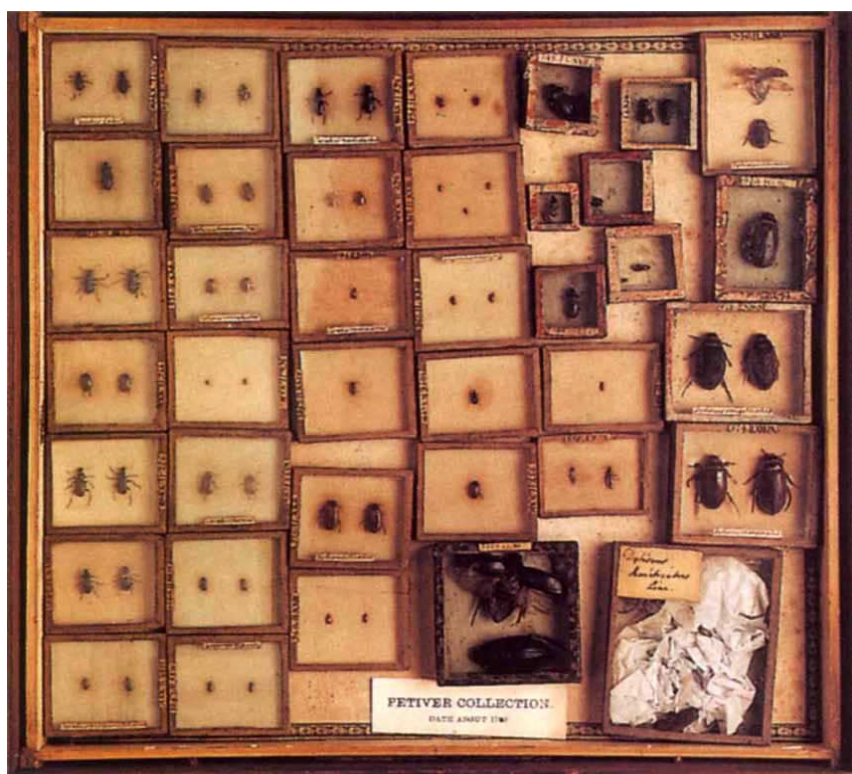


PERSIAN planispheric astrolabe of brass, engraved and inlaid with silver, made for Shah Hussain in 1712 by Abd-al-Ali. From the lavishly illustrated *Sir Hans Sloane* (see over the page).

An opulent physician's "knick-knackatory"



SIR Hans Sloane, the subject of this terracotta bust by Rysbrack, was president both of the Royal College of Physicians and of the Royal Society. His collection of more than 80,000 curiosities and antiques, along with a magnificent library, formed the nucleus of the British Museum after his death in 1753. In *Sir Hans Sloane: Collector, Scientist, Antiquary, Founding Father of the British Museum*, Arthur MacGregor gathers together essays by current and former staff of the British and Natural History Museums. They use the 30 surviving manuscript catalogues of Sloane's original collection to document its contents (which include a late-sixteenth-century watercolour of *English Sailors in a Skirmish with an Eskimo* (left) and a drawer of beetles in small glass-topped bottles (below)) and to consider his interests and aims. Published by British Museum Press in association with Alistair McAlpine, £50.



Highs and lows

Steven W. Van Sciver

Introduction to High-Temperature Superconductivity. By Thomas P. Sheahen. Plenum: 1994. Pp. 580. \$59.50, £47.

THE discovery in 1986 of high-temperature superconductors has led to a rapid growth in research and development in this class of materials, and with it an expansion in the associated literature. Attendance at conferences has greatly increased, new journals have been launched and a variety of monographs have been published. This book is a typical example of the last category. According to the author, it evolved from a series of reports produced by Argonne National Laboratory in Illinois under contract from the Electric Power Research Institute. Its stated purpose is to educate engineers and scientists about the topics that fall under the broad heading of 'high-temperature superconductivity'.

This is not a textbook; it contains few examples and no problems. Rather, it summarizes a wide range of topics from the engineering, physics and materials science of superconductivity to its applications. Each chapter is designed to be self-contained and the text is generally accessible.

Unfortunately, the book falls short of its ambitious goal. In an effort to cover the entire field, the author either presents the subject matter at too elementary a level or resorts to a chronological narrative style more valuable for its historical perspective than as an educational tool. Although the sections on the properties and processing of high-temperature superconductivity provide good overviews of the current state of development, the more fundamental material on the physics of high-temperature superconductivity is not presented in enough depth. Sections on applications, which fill just under a third of the book, are fairly superficial, often referring to specific projects of interest to the author, and will soon be out of date. Furthermore, the title of the book is misleading. Its emphasis is in fact on applications in the electric-power industry; fundamental properties of high-temperature superconductivity are only briefly mentioned and superconducting electronics is all but ignored.

Nevertheless, the book is relatively easy to read and well referenced. It should be a useful source for all workers in the field as well as for nonexperts who want to learn more about this fascinating branch of science and technology. □

Steven W. Van Sciver is at the National High Magnetic Field Laboratory, FAMU/FSU College of Engineering, 1800 E. Paul Dirac Drive, Tallahassee, Florida 32306-4005, USA.

Nanotribology: friction, wear and lubrication at the atomic scale

Bharat Bhushan, Jacob N. Israelachvili & Uzi Landman

Friction, wear and lubrication between materials in contact are of fundamental importance in many pure and applied sciences. Owing to the development of experimental and computer-simulation techniques for studying these phenomena at the atomic scale, an understanding is beginning to emerge of the molecular mechanisms of tribology in thin films and at surfaces.

UNDERSTANDING the atomic processes occurring at the interface of two materials when they are brought together, separated or moved with respect to one another is central to many technological problems, including adhesion, contact formation, friction, wear, lubrication, nanoindentation, fracture and machining¹⁻⁴. At most interfaces of technological relevance, contact occurs at numerous asperities^{1,2}. Consequently, the importance of investigating single asperity contacts in studies of the fundamental micromechanical and tribological properties of surfaces and interfaces has been long recognized⁵⁻¹⁴. The recent emergence and proliferation of proximal probes, in particular tip-based microscopies (the scanning tunnelling microscope and the atomic-force microscope) and the surface-force apparatus, and of computational techniques for simulating tip-surface interactions and interfacial properties, has allowed systematic investigations of interfacial problems with high resolution as well as ways and means for modifying and manipulating nanoscale structures. These advances provide the impetus for research aimed at developing a fundamental understanding of the nature and consequences of the interactions between materials on the atomic scale, and they guide the rational design of materials for technological applications. In short, they have led to the appearance of the new field of nanotribology^{15,16}, which pertains to experimental and theoretical investigations of interfacial processes on scales ranging from the atomic and molecular to the microscale.

The surface-force apparatus (SFA)¹⁷⁻²⁴, the scanning tunnelling microscope (STM)²⁵, the atomic-force and friction-force microscopes (AFM and FFM)²⁶⁻²⁸, as well as the quartz microbalance technique²⁹, are widely used in nanotribological studies. The SFA was developed in the late 1960s and is commonly employed to study both static and dynamic properties of molecularly thin films sandwiched between two molecularly smooth surfaces. The STM, developed in 1981, allows imaging of electrically conducting surfaces with atomic resolution, and has been used for imaging of clean surfaces as well as of lubricant molecules. The introduction of the atomic-force microscope in 1985 provided a method for measuring ultra-small forces between a probe tip and an electrically conducting or insulating surface, and has been used for topographical measurements of surfaces on the nanoscale, as well as for adhesion and electrostatic force measurements. Subsequent modifications of the AFM led to the development of the friction-force microscope (FFM), designed for atomic-scale and microscale studies of friction. This instrument measures forces transverse to the surface. The AFM is also being used for investigations of wear, indentation, detection of transfer of material, boundary lubrication, and nanofabrication and machining. Meanwhile, significant progress in understanding the fundamental nature of bonding and interactions in materials, combined with advances in computer-based modelling and simulation methods, have allowed theoretical studies of complex interfacial phenomena with high resolution in space and time^{15,16}. Such simulations provide insights into atomic-scale energetics, structure, dynamics, thermodynamics, transport and rheological aspects of tribological processes.

Furthermore, these theoretical approaches guide the interpretation of experimental data and the design of new experiments, provide a framework for analysis and development of unifying concepts, and enable the prediction of new phenomena based on atomistic principles.

Here we review several of the principal theoretical and experimental aspects of research in nanotribology. The nature of interactions between two surfaces brought close together, and those between two surfaces in contact as they are separated, have been studied experimentally with the surface-force apparatus. This has led to a basic understanding of the normal forces between surfaces, and of the dramatic way in which these are modified by the presence of an intervening medium such as a liquid or a polymer 'brush' adsorbed or chemically grafted to the surfaces. The frictional properties of such systems have been studied by moving the surfaces laterally, and such experiments have provided insights into the molecular-scale operation of lubricants such as thin liquid or polymer films. Often, these lubricants behave in ways that would not be predicted on the basis of their bulk-scale behaviour alone.

Complementary to these studies are those in which the AFM and FFM are used as a 'model asperity' in contact with a solid surface. These experiments have demonstrated that the relationship between friction and surface topography is not always simple or obvious. AFM studies have also revealed much about the nanoscale nature of intimate contact—of indentation and wear. Computer simulations of tip-surface interactions at the atomic scale have shown the surface need not behave as an elastic medium but can deform plastically, for example leading to 'necking' between surface layers and the probe tip.

Surface-force studies of thin films

The surface-force apparatus has long been used to measure both static^{17,19} and dynamic^{18,20-24} forces between smooth, macroscopic surfaces. As with AFMs and most other force-measuring devices, forces are determined from the deflection of a spring, and displacements are measured using some optical, electrical, capacitive or strain-gauge technique. The principle of operation of SFAs (Fig. 1) is thus fairly simple; the challenge comes in measuring very weak forces, sub-microscopic surface geometries and surface separations at the ångström level.

Static (equilibrium) properties of thin liquid films. The intermolecular interaction potential or force function between two surfaces interacting in or across a liquid medium can be attractive, repulsive, 'oscillatory' or a more complex function of separation. Recent experiments and theoretical studies (mainly computer simulations) have shown that the structure of the liquid molecules and of the confining surfaces are very important in determining what the force function will be³⁰⁻⁴⁵. Furthermore, two crystalline surfaces may induce a liquid to solidify (freeze) when the gap thickness is close to a small multiple of the molecular diameter (Fig. 1, bottom left)^{37-40,46,47}. The oscillatory force-function resulting from such an interaction is illustrated in Fig. 1 (top left)³¹⁻³³.

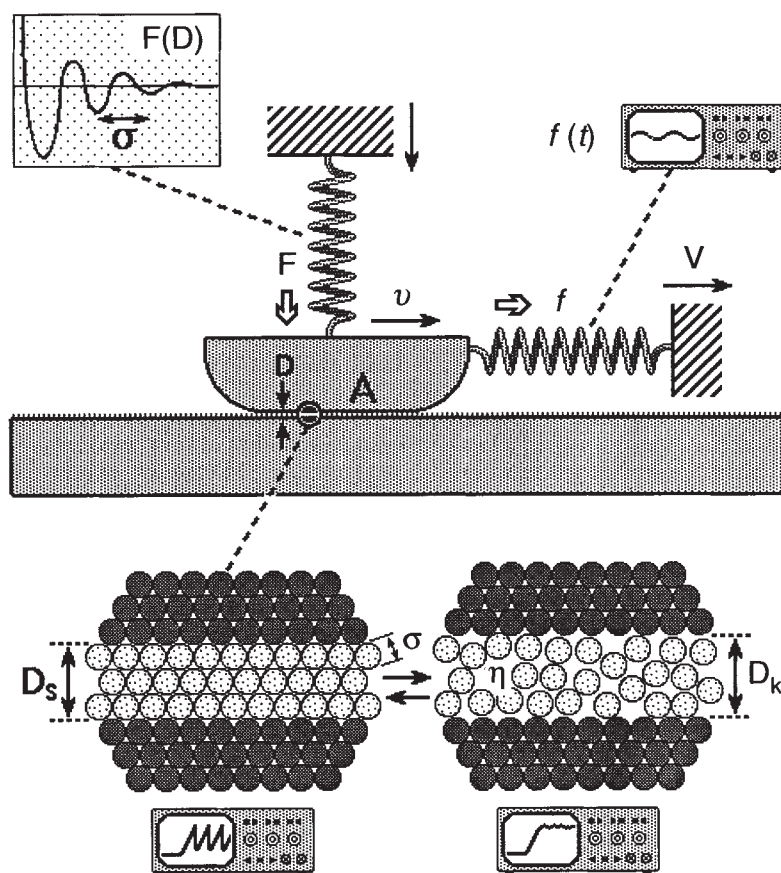


FIG. 1 Illustration of the principle of the surface-force apparatus (SFA) for measuring the normal forces, F , and friction or shear forces, f , between two smooth surfaces of area A separated by a thin liquid film of thickness D (refs 18–24). An optical interference technique allows the shapes of the surfaces and their separation, D , to be accurately monitored during force and friction measurements. Typical contact area A is $100 \mu\text{m}^2$, and gap thicknesses can range from one molecular layer ($D = \sigma = 4 \text{ \AA}$, where σ is the molecular diameter) to thousands of ångströms. An oscillatory force-function, $F(D)$, typically measured across simple liquids is shown at top left, and two common types of friction traces (friction force versus time: $f(t)$) are shown in the oscilloscopes at the bottom of the figure. During sliding, two smooth surfaces usually remain in one of the adhesive minima of the force function, although the kinetic film thickness D_k is usually slightly different from the equilibrium or static thickness, D_s .

The structure adopted by molecules in ultra-thin films is generally very different from their bulk structure; other thin-film properties may also be very different. The freezing point may be higher or lower, depending on whether the 'epitaxial' interaction of the confined molecules with the two surfaces is cooperative or disruptive^{37–40,48} and molecular relaxation times can be many orders of magnitude longer than in the bulk^{23,24,47,49–51}. It is also interesting to note that even though one surface can have an effect on the structure of the liquid molecules adjacent to it^{33,52–55}, the effect of two confining surfaces is generally much more dramatic, and can produce a greater variety of interfacial (non-bulk-like) structures—amorphous, solid or liquid-crystalline—each of which gives rise to a different type of force function as well as to different dynamic properties, as described below.

Dynamic and shear properties of thin films. The force between two surfaces across a liquid or lubricant fluid tells us whether two surfaces, when pressed together, will come into adhesive contact or remain separated by a thin, liquid-like layer^{30–33,36,56}. This is crucially important for understanding the behaviour that results when one of the surfaces is made to move laterally or to shear past the other, as occurs during frictional and lubricated sliding. For example, if the two surfaces are in true molecular contact when sliding begins, their friction will be high and their adhesion may be strong enough to tear them apart. If, however, the surfaces are separated by one or more molecular layers of the fluid, and if this cushioning film remains in the liquid-like state during sliding, this may ensure that the friction force f will be low and that sliding will proceed smoothly (Fig. 1, bottom right). Experiments have shown that even one layer of water molecules, approximately 0.25 nm thick, between two hydrophilic surfaces can be sufficient to reduce the interfacial friction force to a very low value⁵⁷.

On the other hand, if the surface-fluid interactions induce the liquid molecules to solidify, the molecular configuration during sliding will be much more complicated. In such cases the film alternately melts and freezes during the motion^{37–40,46,47,58–60}, as shown schematically in the lower part of Fig. 1; the resulting

friction is of the 'stick-slip' type. Sticking occurs in the frozen state, giving rise to the static friction force, f_s , and slipping occurs in the shear-induced molten state, giving rise to the kinetic friction force, f_k . The real situation, even for the simplest surfaces and liquids, is actually much more complicated than shown in Fig. 1: the direction of sliding is generally at some angle to the surface lattices, the two surface lattices are not generally in registry, the surface and liquid molecules have different diameters, and they are generally not even spherical. This complexity leads to complex force functions and friction-time relationships. Even for quite complex force functions, however, recent experiments have shown that the forces of friction and adhesion are closely related⁶¹.

Shear-induced ordering transitions. Recent experiments and computer simulations have shown that different types of ordered structures may occur during frictional sliding^{37–41,62}. The idea of a 'structure' being induced by motion is an unfamiliar concept, but 'shear-induced ordering' is at the heart of what goes on in such thin films. Different types of ordered states can appear depending on the temperature, load and sliding velocity, and dynamic phase transitions between these states can also occur during sliding, each giving rise to a complex stick-slip friction pattern. Such effects are now believed to underlie many other apparently unrelated dynamic processes involving stick-slip motion, such as earthquakes, the sound of a violin (and other instruments), the squeaking of doors, and sensory perception (of touch and even taste)⁴⁸.

Figure 2 shows some of the dynamic film structures that can arise during the sliding of two surfaces that contain boundary layers of surfactant molecules. In general, solid-like films exhibit stick-slip motion, amorphous-like films exhibit high friction due to the molecular interdigitations and entanglements occurring between the two surfaces, and liquid-like films exhibit low, viscous-like friction with smooth sliding. These dynamic states are not intrinsic properties of a shearing film (in the way that they are for bulk condensed phases at any given temperature and pressure): they depend not only on the temperature and pressure

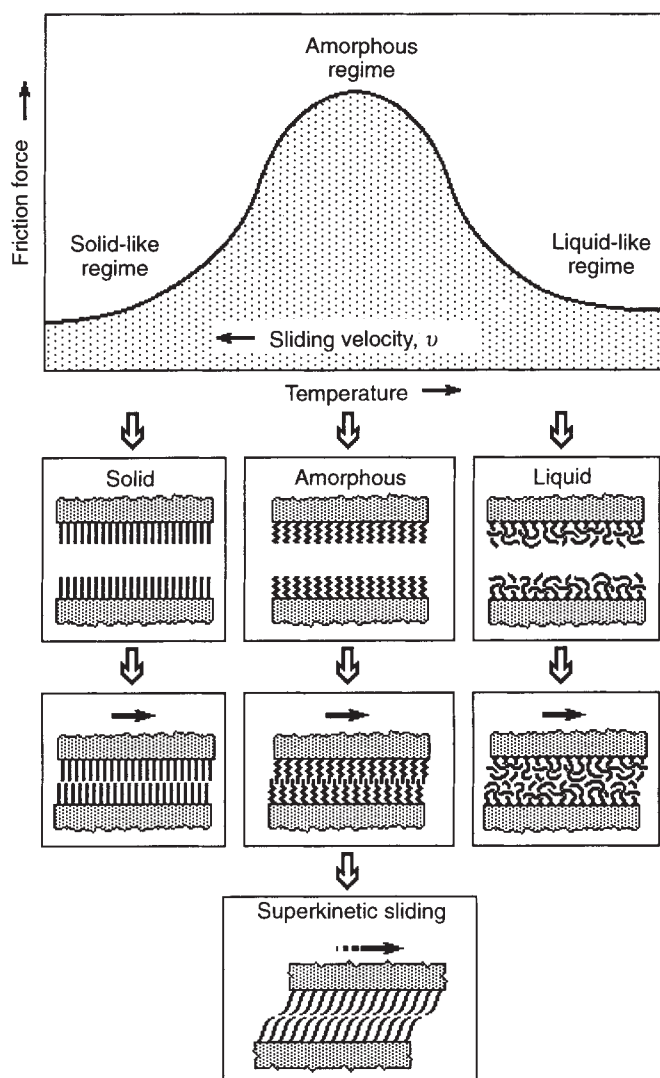


FIG. 2 Schematic illustrations of different types of dynamic thin-film structures that can arise during lubricated sliding. Top, dynamic phase diagram representation of friction force (or rate of energy dissipation) as a function of temperature or sliding velocity (adapted from refs 61 and 62). Under certain favourable conditions involving surface-grafted chain molecules⁶², the entangled molecules at the interface may undergo 'shear-induced ordering', when they disentangle and become aligned; at the same time the friction force drops significantly (superkinetic friction).

or applied load, but also on the relative sliding velocity of the two shearing surfaces (which is not necessarily the same at the velocity of the drive: see Fig. 1). Higher sliding speeds make the film more solid-like, whereas slower speeds make it behave more like a liquid. This principle, known as time-temperature superposition, has recently been found to be very useful in understanding the effects of time, temperature, load and other variables on the complex molecular processes occurring at a shearing interface and the resulting friction forces⁶¹.

Attaining desirable friction properties. The attainment of low adhesion, low friction and low wear (surface damage) is usually a desirable practical goal. But there are cases where the aim may be to attain a constant friction force (with load or speed) or even a high friction, as occurs in the case of clutch fluids. By judiciously choosing the right lubricant fluid and, in some cases, by chemically grafting chain-like molecules such as surfactants⁶² or polymers⁶³ to surfaces (these are known as boundary lubricants), it is becoming possible to control the frictional properties of moving components in machines and other devices. For

example, the grafting of certain surfactant molecules at a certain coverage to surfaces has been found⁶² to produce films that can undergo a dynamic first-order transition from a high friction state to a low friction state (Fig. 2). The advantage of such a lubricant system is that no lubricant fluid is actually ever applied or needed—the surfaces are self-lubricating in the sense that they are processed right from the start to carry their own lubricant layer with them; also, very small quantities of material are needed so that material costs should not be prohibitive even for fairly expensive lubricants. This is a good example of how recent research in the field of molecular tribology is opening up possibilities for creating new types of lubricant systems.

AFM/FFM studies of tribological processes

Adhesion^{64–68}, friction^{27,28,69–77}, wear^{71–73,76–78} and lubrication^{64,79–82} at the interface between two solids with and without liquid films have been studied using the AFM and FFM. Surface roughness is routinely measured using the AFM^{83–86}, and both instruments have been used for measuring elastic-plastic mechanical properties^{72,73,76,77,87–89}. At most solid-solid interfaces of technological relevance, contact occurs at numerous asperities^{1–4,15,16}; a sharp AFM/FFM tip sliding on a surface simulates just one such contact^{15,16}. For measurements of surface roughness, friction forces and nanoscale scratching and wear, a microfabricated square-pyramidal Si_3N_4 or silicon tip with a tip radius ranging from 10 to 50 nm is generally used¹⁶ at loads ranging from 10 to 150 nN. For measurements of microscale scratching and wear, and for nanoindentation hardness measurements and nanofabrication, a three-sided pyramidal single-crystal natural-diamond tip with a tip radius of about 100 nm is generally used¹⁶ at relatively high loads, ranging from 10 to 150 μN .

Surface roughness, adhesion and friction. Solid surfaces, irrespective of their method of formation, generally contain surface irregularities. When two nominally flat surfaces are placed in contact, surface roughness causes contact to occur at discrete contact points. Deformation occurs at these points, and may be either elastic or plastic, depending on the nominal stress, surface roughness and material properties. The sum of the areas of all the contact points constitutes the real area of contact, and for most materials at normal loads, this will be only a small fraction of the area that would be in contact if the surfaces were perfectly smooth. In general, this real area of contact must be minimized to minimize adhesion, friction and wear^{1,2}.

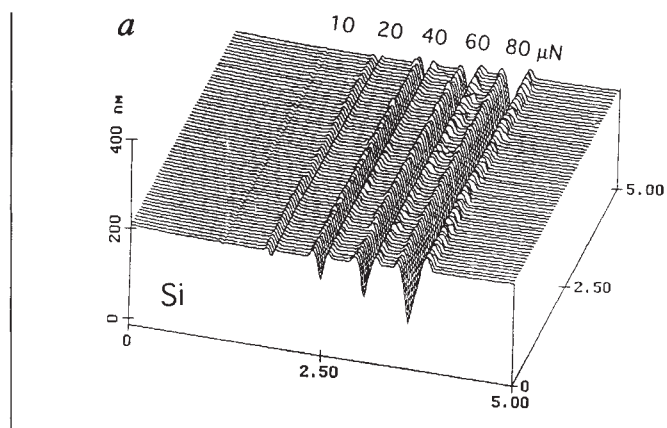
Characterizing surface roughness is therefore important for predicting and understanding the tribological properties of solids in contact. The AFM has been used to measure surface roughness on length scales from nanometres to micrometres^{2,83–86}. Surface roughness most commonly refers to the variations in the height of the surface relative to a reference plane^{1,2}. Commonly measured roughness parameters, such as r.m.s. surface height and peak-to-valley distance, are found to be scale-dependent for any given surface. The topography of most engineering surfaces is fractal, possessing a self-similar structure over a range of scales. By using fractal analysis one can characterize the roughness of such surfaces with two scale-independent fractal parameters D and C , which provide information about roughness at all length scales^{83–86}. These two parameters are instrument-independent and are unique for each surface. D (generally ranging from 1 to 2) primarily relates to distribution of different frequencies in the surface profile, and C to the amplitude of the surface height variations at all frequencies. A fractal model of elastic-plastic contacts⁸⁴ has been used to predict whether contacts experience elastic or plastic deformation, and to predict the statistical distribution of contact points.

To study friction mechanisms on an atomic scale, a well characterized, freshly cleaved surface of highly oriented pyrolytic graphite (HOPG) has been studied by Mate *et al.*²⁷ and Ruan and Bhushan⁷⁴. The atomic-scale friction force of HOPG exhibited the same periodicity as the corresponding topography

(Fig. 3a), but the peaks in friction and those in topography were displaced relative to each other⁷⁴ (Fig. 3b). A Fourier expansion of the interatomic potential was used to calculate the conservative interatomic forces between atoms of the FFM tip and those of the graphite surface. Maxima in the interatomic forces in the normal and lateral directions do not occur at the same location, which explains the observed shift between the peaks in the lateral force and those in the corresponding topography. Furthermore, the observed local variations in friction can be explained by variation in the intrinsic lateral force between the sample and the FFM tip⁷⁴, and these variations may not necessarily occur as a result of atomic-scale stick-slip processes^{27,90,91}.

Frictional forces of HOPG have also been studied on micro-metre scales^{70,75–77,92}. Local variations in the microscale friction of cleaved graphite are observed, which arise from structural changes that occur during the cleaving process⁷⁵. The cleaved HOPG surface is largely atomically smooth but exhibits line-shaped regions in which the coefficient of friction is more than order of magnitude larger. Transmission electron microscopy indicates that the line-shaped regions consist of graphite planes of different orientation, as well as of amorphous carbon. Differences in friction can also be seen for organic mono- and multi-layer films⁷⁰, which again seem to be the result of structural variations in the films. These measurements suggest that the FFM can be used for structural mapping of the surfaces. FFM measurements can be used to map chemical variations, as indicated by the use of the FFM with a modified probe tip to map the spatial arrangement of chemical functional groups in mixed organic monolayer films⁹². Here, sample regions that had stronger interactions with the functionalized probe tip exhibited larger friction.

Local variations in the microscale friction of scratched



surfaces can be significant, and seen to depend on the local surface slope rather than the surface height distribution^{16,71–77}. Directionality in friction is sometimes observed on the macro-scale; on the microscale this is the norm^{16,28,71–77}. This is because most 'engineering' surfaces have asymmetric surface asperities, so that the interaction of the FFM tip with the surface is dependent on the direction of the tip motion. Moreover, during surface finishing processes material can be transferred preferentially onto one side of the asperities, which also causes asymmetry and directional dependence. Reduction in local variations and in directionality of frictional properties therefore requires careful optimization of surface roughness distributions and of surface-finishing processes.

Table 1 shows the coefficients of friction measured for two surfaces on micro- and macroscales. The coefficient of friction

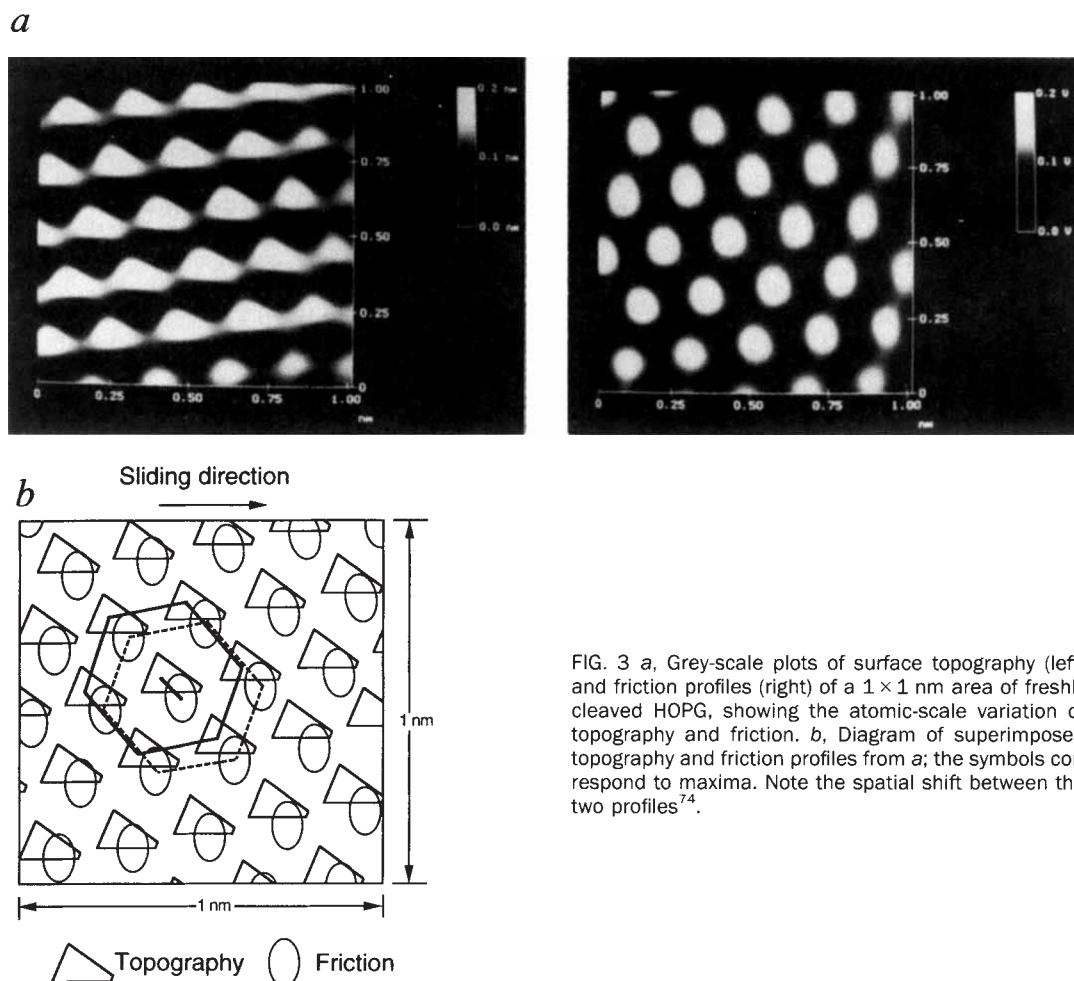


FIG. 3 a, Grey-scale plots of surface topography (left) and friction profiles (right) of a 1×1 nm area of freshly cleaved HOPG, showing the atomic-scale variation of topography and friction. b, Diagram of superimposed topography and friction profiles from a; the symbols correspond to maxima. Note the spatial shift between the two profiles⁷⁴.

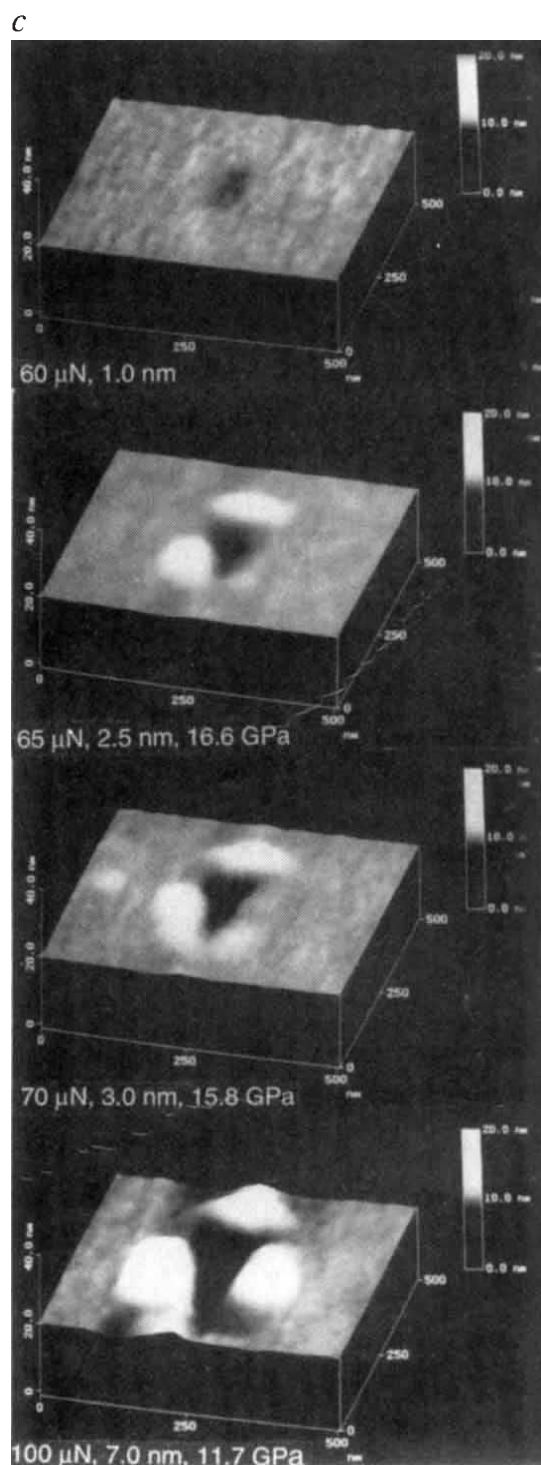
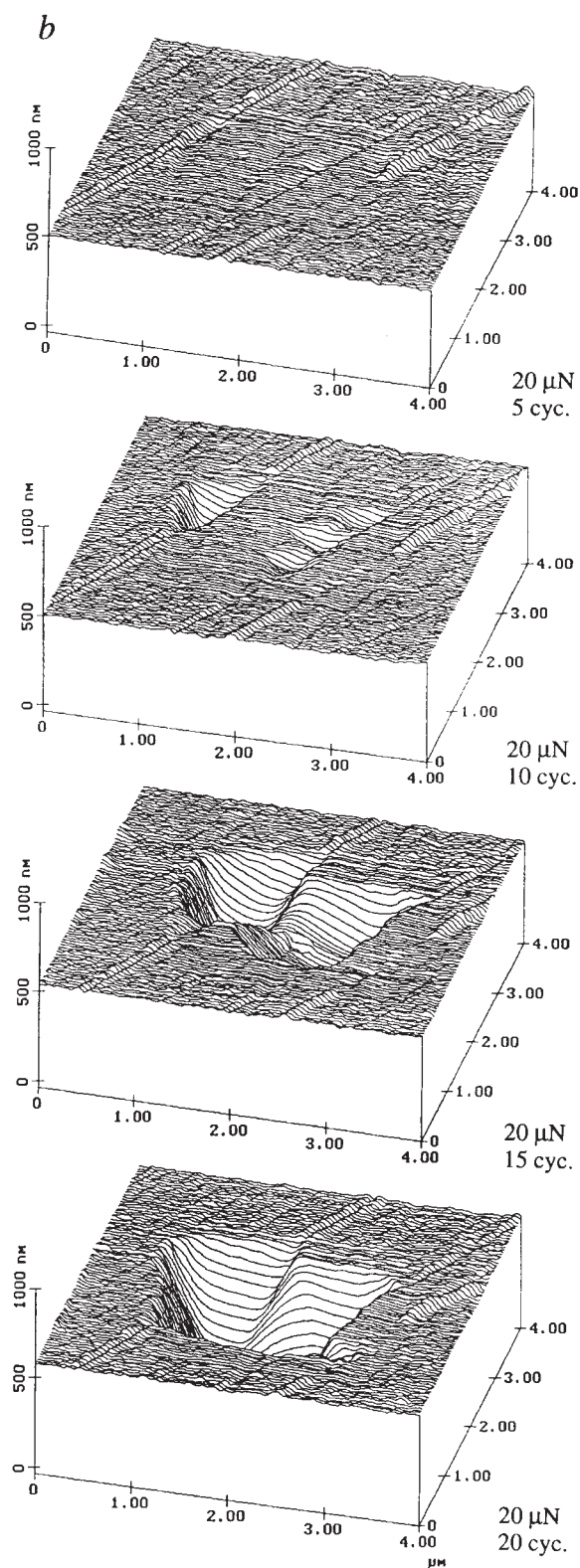


FIG. 4 *a*, Surface profile of Si(111) scratched at various loads⁷³; note that the *x* and *y* axes are in μm , and the *z* axis in nm. *b*, Surface profiles of diamond-like carbon-coated thin-film disk showing the worn region; the normal load and the number of test cycles are indicated⁷². (Units of axes as in *a*.) *c*, Grey-scale plots of images of indentation marks on the Si(111) samples; loads, indentation depths and hardness values are indicated⁸⁹. (All axes in nm.)

is defined as the ratio of the friction force to the normal load. The values on the microscale are much lower than those on the macroscale. When measured for the small contact areas and very low loads used in microscale studies, indentation hardness and modulus of elasticity are higher than at the macroscale. This reduces the degree of wear. In addition, the small apparent areas

of contact reduce the number of particles trapped at the interface, and thus minimize the 'ploughing' contribution to the friction force.

At higher loads, however, the coefficient of friction for microscale measurements increases towards values comparable with those obtained from macroscale measurements, and surface

TABLE 1 Surface roughness and coefficients of friction

Material	r.m.s. roughness (nm)	Micro-scale coefficient of friction versus Si ₃ N ₄ tip*	Macro-scale coefficient of friction versus alumina ball†
Si(111)	0.11	0.03	0.18
C ⁺ -implanted Si	0.33	0.02	0.18

Data from refs 73 and 81.

* Tip radius of ~ 50 nm in the load range 10–150 nN (2.5–6.1 GPa), a scanning speed of $5 \mu\text{m s}^{-1}$ and scan area of $1 \times 1 \mu\text{m}$.

† Ball radius of 3 mm at a normal load of 0.1 N (0.3 GPa) and average sliding speed of 0.8 mm s^{-1} .

damage also increases. Thus Amontons' law of friction^{1,2}, which states that the coefficient of friction is independent of apparent contact area and normal load, does not hold for microscale measurements. These findings suggest that microcomponents sliding under lightly loaded conditions should experience very low friction and near-zero wear.

Scratching, wear and indentation. The AFM can be used to investigate how surface materials may be moved or removed on micro- to nanoscales, for example through scratching and wear^{71, 73, 76–78} (where these things are undesirable), and nanomachining and nanofabrication (where they are desirable). The AFM can also be used for measurements of mechanical properties on micro- to nanoscales. Figure 4a shows microscratches made on Si(111) at various loads after 10 cycles. As expected, the depth of scratch increases with load. Such microscratching measurements can be used to study failure mechanisms on the microscale and to evaluate the mechanical integrity (scratch resistance) of ultra-thin films at low loads^{71, 72}.

By scanning the sample in two dimensions with the AFM, wear scars are generated on the surface. The evolution of wear of a diamond-like carbon coating on a polished aluminium substrate is shown in Fig. 4b, which illustrates how the micro-wear profile for a load of 20 μN develops as a function of the number of scanning cycles⁷². Wear is not uniform, but is initiated at nanoscratches, indicating that surface defects (with high surface energy) act as initiation sites. Thus, scratch-free surfaces will be relatively resistant to wear.

Mechanical properties, such as hardness and modulus of elasticity can be determined on micro- to picoscales using the AFM^{72, 73, 76, 77, 87–89}. Indentability on the scale of picometres can be studied by monitoring the slope of cantilever deflection as a function of sample travelling distance after the tip is engaged

and the sample is pushed against the tip⁷¹. For a rigid sample, the cantilever deflection is the same as the sample travelling distance; but the former quantity is smaller if the tip indents the sample. The indentation hardness of surface films with an indentation depth as small as 1 nm has been measured for Si(111)⁸⁹ (Fig. 4c). Triangular indentations are observed for shallow penetration depths. The hardness for an indentation depth of 2.5 nm is 16.6 GPa; it drops to 11.7 GPa at a depth of 7 nm. This decrease in hardness with increase in indentation depth can be rationalized on the basis that as the volume of deformed material increases, there is a higher probability of encountering material defects⁹³. Bhushan and Koinkar⁷³ have used AFM measurements to show that ion implantation of silicon surfaces increases their hardness and thus their wear resistance. Formation of surface alloy films with improved mechanical properties by ion implantation is of growing technological importance as a means of improving the mechanical properties of materials.

The Young's modulus of elasticity is calculated from the slope of the indentation curve during unloading^{16, 87}. Maivald *et al.*⁸⁸ used an AFM in 'force modulation mode' to measure surface elasticities: the AFM tip is scanned over the modulated sample surface with the feedback loop keeping the average force constant. For the same applied force, a soft area deforms more, and thus causes less cantilever deflection, than a hard area. The ratio of modulation amplitude to the local tip deflection is then used to create a 'force modulation image'. The force modulation mode makes it easier to identify soft areas on hard substrates.

Detection of the transfer of material on a nanoscale is possible with the AFM. Indentation of C₆₀-rich fullerene films with an AFM tip has been shown⁹⁴ to result in the transfer of fullerene molecules to the AFM tip, as indicated by discontinuities in cantilever deflection as a function of sample travelling distance in subsequent indentation studies.

Boundary lubrication. The 'classical' approach to lubrication uses freely supported multimolecular layers of liquid lubricants^{64, 79, 80}. But for lubrication of microdevices, a more effective approach involves the deposition of organized, dense molecular layers of long-chain molecules on the surfaces in contact. Such monolayers and thin films are commonly produced by Langmuir–Blodgett (LB) deposition and by chemical grafting of molecules into self-assembled monolayers (SAMs)⁸¹. All of these approaches involve boundary lubrication. The AFM and FFM have been used to study the lubricating effects of freely supported, LB and SAM films^{81, 82}. For freely supported liquid lubricants, either increasing the film thickness or chemically

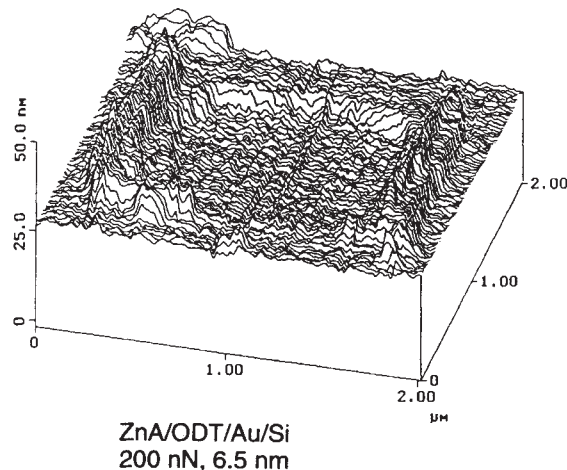
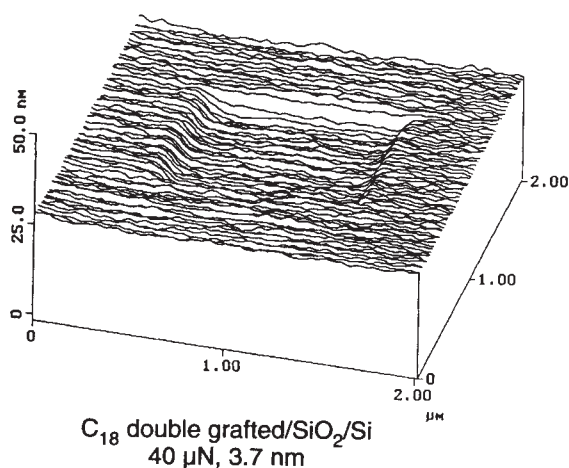


FIG. 5 Surface profiles showing the worn region after one scan cycle for self-assembled monolayers of octadecyl silanol (left) and zinc arachidate (right). Normal force and the wear depths are indicated. Note

that wear or ZnA film occurs at only 200 nN as compared to 40 μN for C₁₈ film⁸¹.

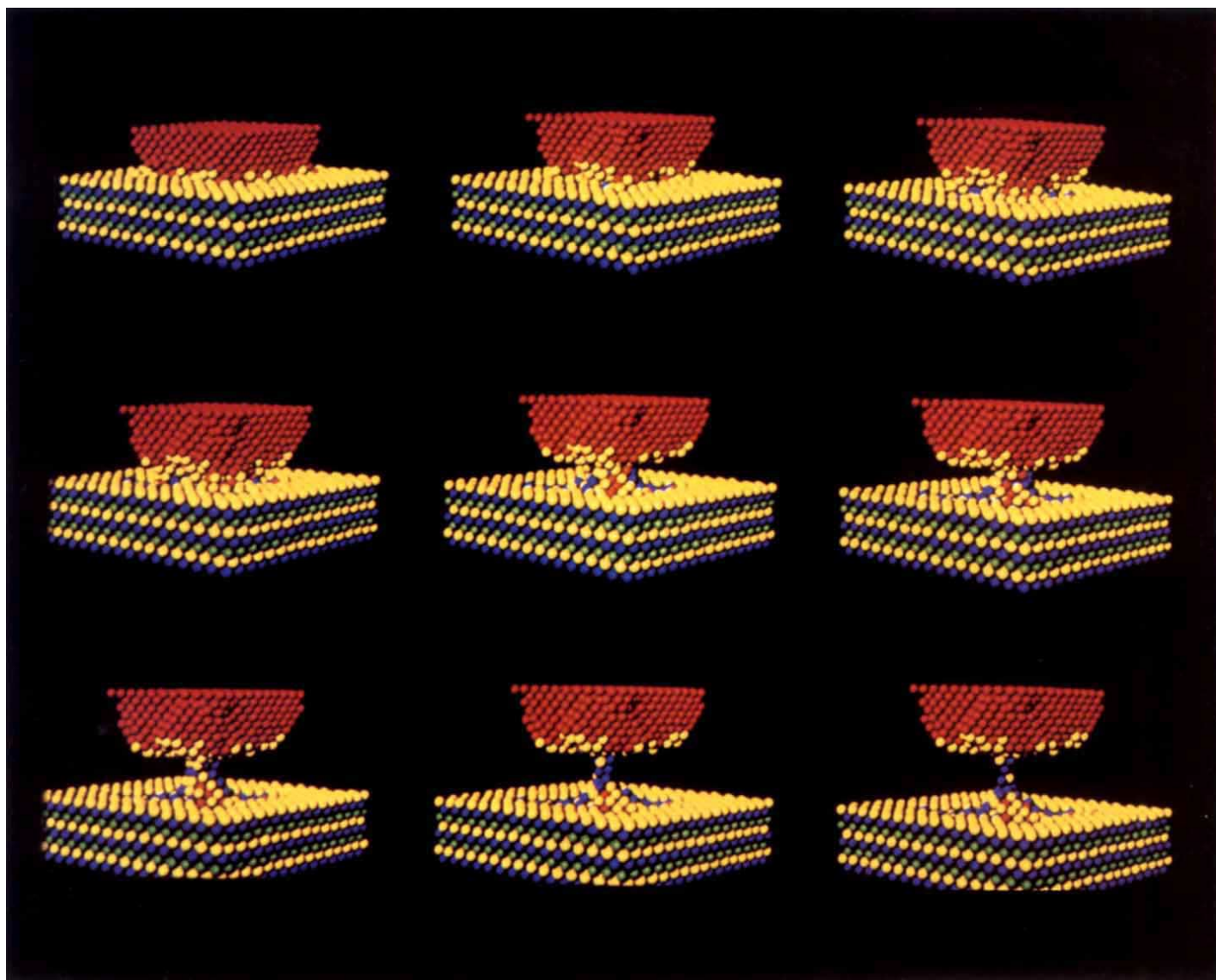


FIG. 6 Sequence of atomic configuration starting from a Ni tip indented in an Au(001) substrate (top left) and during the process of retraction

of the tip (from left to right) accompanied by formation of a connective solid gold junction (bottom right). Red balls represent tip nickel atoms.

bonding the molecules to the substrate improves the lubrication performance⁸². SAMs of octadecyl (C_{18}) compounds based on aminosilanes on an oxidized silicon surface exhibit a lower coefficient of friction (0.018) and greater durability than LB films of zinc arachidate adsorbed on a gold surface coated with octadecylthiol (ODT) (coefficient of friction, 0.03) (Fig. 5)⁸¹. LB films are bonded to the substrate only by weak van der Waals attraction, whereas SAMs are chemically bound via covalent bonds. Because of the choice of chain length and terminal linking group that SAMs offer, they hold great promise for boundary lubrication of microdevices.

Measurements of the thickness of ultra-thin lubricant films with nanometre lateral resolution can be made with the AFM^{95–97}. The lubricant thickness is obtained by measuring the force on the tip as it approaches, contacts and pushes through the liquid film and ultimately contacts the substrate. The distance between the sharp 'snap-in' (owing to the formation of a liquid meniscus between the film and the tip) at the liquid surface and the hard repulsion at the substrate surface is a measure of the liquid film thickness. This technique is now used routinely in the information-storage industry for thickness measurements (with

nanoscale spatial resolution) of lubricant films, a few nanometres thick, in rigid magnetic disks.

Atomic-scale simulations

An understanding of tribological phenomena at the molecular scale requires the non-trivial task of unravelling the energetics, structure, dynamics and transport processes of non-uniform systems involving a whole host of materials (metals, ionic solids, semiconductors, ceramics, organic and polymeric materials), under nonequilibrium conditions and often beyond the linear-response regimes for mechanical and flow behaviour. Consequently, until rather recently, most theoretical approaches to tribological problems have been anchored in continuum formulations, which assume continuous materials properties and use the formalisms of contact mechanics, hydrodynamics, viscous flow and elastohydrodynamics. But advances in theoretical understanding of the nature of interatomic interactions in materials, combined with the development of methods for computer-based modelling of complex many-body systems, have now opened up new avenues for exploring the properties of materials at high spatial and temporal resolution⁹⁸. These studies employ

large-scale molecular-dynamics simulations, in which the trajectories in phase space of systems consisting of thousands of atoms (and subject to appropriate boundary conditions) are calculated from the particles' newtonian equations of motion. Generally the nature of the many-body interactions are derived from quantum-mechanical calculations. Analyses of particle trajectories allow the determination of structural, energetic, dynamical and mechanical properties of the system on nanometre distance scales and femtosecond timescales. Some of the results obtained from such simulations correlate well with the predictions of theories constructed on the basis of macroscopic considerations^{35,99}. Thus these results help to elucidate the microscopic origin of the macroscopic behaviour.

On the other hand, as we have seen to be the case experimentally, simulations often reveal that the physical behaviour of materials at interfaces can be very different from that in the bulk. These interface-specific phenomena include atomic-scale adhesion and interfacial wetting on contact formation^{66,99}, generation of atomic-scale connective wires by the separation of metal-metal contacts^{66,100}, atomic-scale stick-slip friction^{56,90,91} and structural ordering and phase transitions in thin lubricant films pressed and sheared between confining solid plates^{46,58,60,101}. Some of these effects are quite general to systems confined to molecular dimensions in one or more directions, and they may determine the range of validity of macroscopic descriptions of the mechanical, rheological and flow properties of materials. Atomic-scale simulations can be of great value for interpreting the results of micro- and nanotribological experiments, particularly those involving scanning-probe microscopies, the surface-force apparatus and microbalance techniques. Moreover, understanding the behaviour of materials of very small dimensions is relevant to the development of nano- and microfabricated devices and to the atomic-scale manipulation of materials^{98,102}.

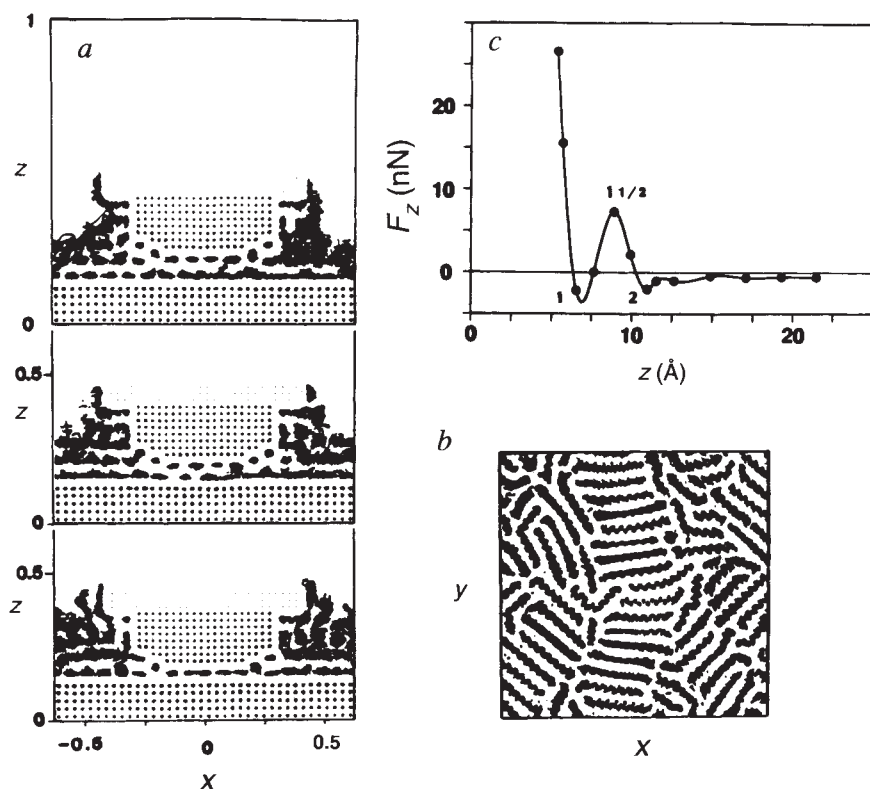
Interfacial solid junctions. Attempts to understand the key issue in friction—the origins and nature of excitations, and the transfer and dissipation of energy occurring when two bodies slide in contact—date back several centuries¹⁰³. For metals, the current view (which emerged over four decades ago) is that the friction force is essentially the force required to shear intermetallic

junctions formed at the regions of real contact, plus the force required to plough the surface of the softer material by the asperities of the harder¹.

Simulations of a clean gold substrate contacted by a nickel tapered and faceted tip, and for the reverse situation of a nickel surface and a gold tip, revealed^{66,100} the onset of an instability as the tip approaches the sample to a distance of about 4 Å. At this point there occurs a jump to contact (with gold atoms being displaced by about 2 Å in about 1 picosecond), with adhesive bonding between the two materials driven and accompanied by atomic-scale wetting of nickel by gold atoms. The latter is the result of differences in their surface energies, just as it is for the case of surface wetting by a liquid film. Retraction of the tip from the surface after contact causes significant inelastic deformation of the sample, involving ductile extension, formation of a connective neck of atomic dimensions, and eventual rupture. The final result on separation is a gold-coated nickel tip and a damaged gold surface. These phenomena, which were observed also in simulations in which the tip penetrated the surface, involve transitions from elastic to plastic behaviour of the metals. The formation of an extended, atomically thin crystalline connective wire (Fig. 6) gives rise to marked hysteresis in the force-distance relation between the materials on lowering and lifting the tip. The elongation of the crystalline junction exhibits straining and yielding stages, with a periodicity equal to that of the interlayer spacing in the wire, which is reflected in periodic oscillations in the recorded force. This behaviour has been seen experimentally in AFM studies of solid junctions⁶⁶, and also in room temperature STM measurements of the electrical conductance of atomically thin gold wires up to 50 Å long^{104,105} formed between the tip and a gold surface, exhibit a periodic quantization in units of the conductance quantum ($2e^2/h$, where e is the electron charge and h is Planck's constant). For longer wires, of lengths greater than 100 Å, the onset of localization was observed¹⁰.

Simulations of the interactions between the surfaces of crystalline ionic solids (CaF_2)⁵⁶ and between semiconductor surfaces (silicon)^{90,91} exhibit similar jumps to contact and force hysteresis. Ionic solids are less easily deformed plastically, however; rather,

FIG. 7 a, Side views of atomic configurations of a nickel tip, an Au(001) surface and an $n\text{-C}_{16}\text{H}_{34}$ (hexadecane) film, starting at a tip-to-surface spacing corresponding to two layers of the confined film ($d_{ts} = 11$ Å, top) and ending with a one-layer film ($d_{ts} = 6.5$ Å, bottom). The middle configuration corresponds to spacing $d_{ts} = 9$ Å. The linear dimension of the system along the x axis is 77.5 Å, and the configurations were obtained in a vertical slice in the xz plane of a width of 16 Å. In this figure, the coordinate axes are normalized such that $x=z=1$ corresponds to a linear dimension of 77.5 Å. b, Short-time trajectories of hexadecane molecules in the layer next to the Au(001) surface, corresponding to the one-layer confined film shown at the bottom of a, illustrating intra-layer ordered domains. (The linear dimension of the system along the x and y axes is 77.5 Å.) c, Normal force on the nickel tip versus the distance between the tip and the gold surface, z , recorded during the process of lowering of the nickel tip onto the hexadecane-covered Au(001) surface, and illustrating force oscillations associated with the layering transformations in the film.



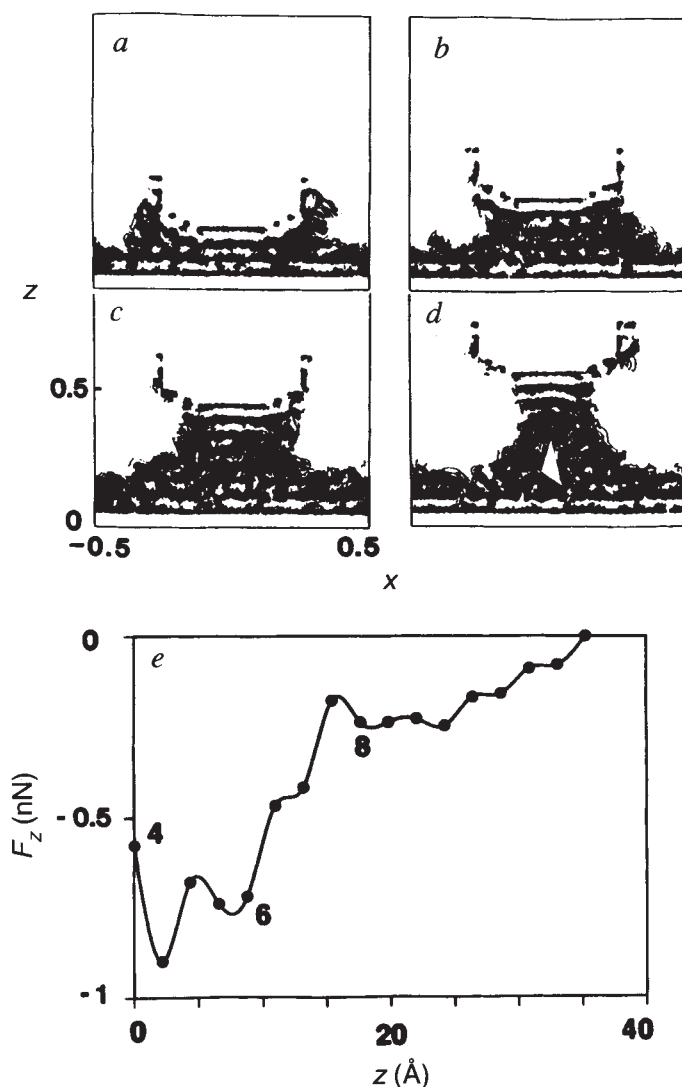


FIG. 8 a–d, Short-time trajectories, obtained from molecular-dynamics simulations, of hexadecane molecules at 350 K forming a capillary column between a gold substrate and a nickel tip at four stages of the tip-lifting process. Trajectories were plotted in a slice (23 Å wide) through the middle of the system. The distances (Å) between the tip and the substrate for the configurations shown in a–d are $d_{ts} = 19.3$, 28.1, 36.9 and 45.7 Å, respectively. The length scale of the calculational cell, which was periodically repeated along the x and y directions, is 77.5 Å. (In this figure, the coordinate axes are normalized such that $x = z = 1$ corresponds to a linear dimension of 77.5 Å.) e, The recorded normal force on the nickel tip (F_z) versus the distance between the tip and the gold surface, z , with the origin taken as that corresponding to a four-layer capillary column (see a). Dots indicate distances for which the system was relaxed during the tip lifting process. The numbers on the graph denote the number of layers in the capillary column (for the eight-layer column, the number is ambiguous due to disordering in the middle part, see c). The changes in the structural nature of the film on elongation of the capillary junction are reflected in the variation in F_z .

they tend to exhibit brittle failure. Such simulations of two materials in contact sliding with respect to one another revealed atomic-scale stick-slip behaviour (oscillatory bond forming and bond breaking)^{90,91}, shearing of the probe tip and interfacial material transfer⁵⁶, and fragmentation owing to asperity collisions¹⁰⁶. Some of these results have been used to calculate the critical yield stress of sheared ionic interfaces and its temperature dependence; these predictions can again be tested against experiments.

Simulations of atomic-scale friction between diamond (111) surfaces (chosen because of the potential use of diamond as a

low-friction wear-resistant coating) have revealed¹⁰⁷ the atomic-scale mechanisms of energy transfer between fully hydrogen- or methyl-terminated diamond surfaces. Energy transfer is relatively efficient, and thus the friction coefficients are relatively high. But the friction coefficient is reduced when longer hydrocarbons are chemisorbed to the surface. For low loads, this reduction can be ascribed to alignment of the molecules in the sliding direction; at high loads it is the result of diminished corrugation of the underlying surface, caused by the tendency of the chemisorbed molecules to occupy corrugation troughs.

Interfacial liquid junctions and confined films. Molecular-dynamics simulations of the nanotribology of interfacial liquids and confined films aim to explore the relation between the physical properties and response of model lubricants and their molecular characteristics, such as chain lengths and molecular structural complexity (for example, straight versus branched chains). Such simulations of alkanes and self-assembled boundary-lubricant systems have revealed liquid-junction formation^{35,36}, inhibition of the generation of solid junctions⁵⁶ layering phenomena and in-plane ordering owing to confinement^{34,45} and have shown how self-assembled monolayers^{42,108–110} respond to compression by a tip or flat surface. Simulations have also been performed of collapse and drainage mechanisms of confined molecular films and their dependence on the strength of bonding to the substrate³⁴, the dependence of thin-film viscosities on the shear rate¹¹¹ and the origin of stick-slip phenomena in shear-induced ‘solidification’ and melting^{46,60,101}, mechanisms of cavitation in thin molecular liquid films induced by fast motion of immersed tips¹¹², and correlations between structural and dynamical transformations in lubricating films with force-distance characteristics^{35,36,112} (Figs 7 and 8). Simulations of thin-film model lubricants made of mixtures of long-chain molecules of different lengths have illuminated the molecular mechanism of preferential surface segregation of the longer-chain molecules⁵⁵.

These simulations have generally considered flat, atomically structured crystalline solids as the substrates. Some recent simulations, on the other hand, have modelled interactions between surface asperities for a sliding metallic contact lubricated by a thin (~ 20 Å) film of hexadecane¹¹³. These simulations are relevant to many technological problems, ranging from elastohydrodynamic lubrication of gears, rolling bearings and cam-followers, to tribology in future generations of information-storage devices^{2,4}. For relative sliding velocities in the range 10 – 20 m s^{−1}, spatial and temporal variations in the density of the lubricant, particularly in the region between the approaching asperities can lead to glassification and seizure of the contact. In addition, sliding asperity contacts can induce dynamical structural layering of the lubricant resulting in oscillatory behaviour of the friction forces and ultimately breakdown of the lubricant film, up to the molecular level, accompanied by trapping of molecules in the space between the two asperities. The asperities themselves can be deformed markedly, leading to the formation of metal-metal junctions and transfer of material between the surfaces. Cavitation was observed in the region between two departing asperities following collision. Such simulations extend investigations of elastohydrodynamic lubrication to the nano-scale domain, which is beyond the range of applicability of continuum models^{114–116}.

Concluding remarks

These theoretical and experimental studies of tribology at the molecular scale have not just extended, but in some cases fundamentally altered, the picture that had developed from macroscale studies. For example, the traditional picture of stick-slip motion ascribed it to a negative slope in the friction-velocity function^{1,2}, but recent work has put this in reverse: in many cases stick-slip behaviour is seen to be the primary process—arising from successive freezing and melting transitions of the shearing film—and it is this that gives rise to the negative friction-velocity

dependence⁴⁶. SFA experiments on films of complex fluids and polymers have also shown why branched-chain molecules are better lubricants than straight-chain molecules, even though the former have much higher bulk viscosities: the symmetrically shaped straight-chain molecules are prone to ordering and freezing, which dramatically increases their resistance to shear, whereas the irregularly shaped, branched molecules remain in the liquid state even under high loads. Straight-chain molecules, however, make ideal clutch fluids, for which high friction is required.

Investigations of wear, scratching and indentation on nanoscales using the AFM can provide insights into failure mechanisms of materials. Coefficients of friction, wear rates and mechanical properties such as hardness have been found to be different on the nanoscale than on the macroscale: generally, coefficients of friction and wear rates on micro- and nanoscales are smaller, whereas hardness is greater.

As the trend towards miniaturization of technological devices continues, these studies indicate that very small micromechanical machines may behave in ways that cannot be predicted from their larger counterparts. It is encouraging in this regard to find that materials properties at small scales can be superior—wear is lower, friction less. Moreover, new lubrication strategies such as the use of self-assembled monolayers promise to be very versatile and effective at these scales. On the other hand, phenomena such as shear-induced solidification and the formation of bridges between contact asperities indicate that the nanotribological regime poses novel challenges of its own. □

B. Bhushan is at the Computer Microtribology and Contamination Laboratory, 206 West 18th Avenue, The Ohio State University, Columbus, Ohio 43210-1107, USA. J. N. Israelachvili is at the Department of Chemical and Nuclear Engineering, University of California, Santa Barbara, California 93106, USA. U. Landman is at the School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

- Bowden, F. P. & Tabor, D. *The Friction and Lubrication of Solids* Parts I & II (Clarendon, Oxford, 1950 & 1964).
- Bhushan, B. *Tribology and Mechanics of Magnetic Storage Devices* (Springer, New York, 1990).
- Bhushan, B. & Gupta, B. K. *Handbook of Tribology: Materials, Coatings and Surface Treatments* (McGraw-Hill, New York, 1991).
- Bhushan, B. *Mechanics and Reliability of Flexible Magnetic Media* (Springer, New York, 1992).
- Gane, N. & Bowden, F. P. *J. appl. Phys.* **39**, 1432–1435 (1968).
- Skinner, J. & Gane, N. *J. Phys.* **D8**, 2087–2094 (1972).
- Pollock, H. M., Shuffelbottom, P. & Skinner, J. *J. Phys.* **D10**, 127–138 (1977).
- Maugis, D., Desatos-Andarelli, G., Heurtel, A. & Courtel, R. *Am. Soc. Lubrication Engrs* **21**, 1–19 (1978).
- Pollock, H. M. *J. Phys.* **D11**, 39–54 (1978).
- Pashley, M. D., Pethica, J. B. & Tabor, D. *Wear* **100**, 7–31 (1984).
- Maugis, D. & Pollock, H. M. *Acta metall.* **32**, 1323–1334 (1984).
- Pethica, J. B. *Phys. Rev. Lett.* **57**, 3235 (1986).
- Pethica, J. B. & Oliver, W. C. *Physica Scripta* **19**, 61–66 (1987).
- Guo, Q., Ross, J. D. J. & Pollock, H. M. *Mater. Res. Soc. Proc.* **140**, 51–66 (1989).
- Fundamentals of Friction: Macroscopic and Microscopic Processes* (eds Singer, J. L. & Pollock, H. M.) (Kluwer, Dordrecht, 1991).
- Handbook of Micro/Nanotribology* (ed. Bhushan, B.) (CRC, Boca Raton, Florida, 1995).
- Tabor, D. & Winterton, R. *Proc. R. Soc. A* **312**, 435–450 (1969).
- Israelachvili, J. N. & Tabor, D. *Nature* **241**, 148–149 (1973); *Wear* **24**, 386–390 (1973).
- Israelachvili, J. N. *Chemtracts—Analyt. phys. Chem.* **1**, 1–12 (1989).
- Briscoe, B. J., Evans, D. C. & Tabor, D. *J. Colloid Interface Sci.* **61**, 9–13 (1977).
- Israelachvili, J. N., Homola, A. M. & McGuiggan, P. M. *Science* **240**, 189–191 (1988).
- Hirz, S. J., Homola, A. M., Hadzioannou, G. & Franc, C. W. *Langmuir* **8**, 328–333 (1992).
- Granick, S. *Science* **253**, 1374–1379 (1991).
- Klein, J., Perahia, D. & Warburg, S. *Nature* **352**, 143–145 (1991).
- Binnig, G., Rohrer, H., Gerber, Ch. & Weibel, E. *Phys. Rev. Lett.* **49**, 57–61 (1982).
- Binnig, G., Quate, C. F. & Gerber, Ch. *Phys. Rev. Lett.* **56**, 930–933 (1986).
- Mate, C. M., McClelland, G. M., Erlandsson, R. & Chiang, S. *Phys. Rev. Lett.* **59**, 1942–1945 (1987).
- Meyer, G. & Amer, N. M. *Appl. Phys. Lett.* **57**, 2089–2091 (1990).
- Krim, J., Solina, D. H. & Chiarello, R. *Phys. Rev. Lett.* **66**, 181–184 (1991).
- Christenson, H. K. *Chem. Phys. Lett.* **118**, 455–458 (1985); *J. phys. Chem.* **90**, 4–6 (1986).
- Gee, M. L. & Israelachvili, J. N. *J. chem. Soc., Faraday Trans.* **86**, 4049–4058 (1990).
- McGuiggan, P. M. & Israelachvili, J. N. *J. Mater. Res.* **5**, 2232–2243 (1990).
- Israelachvili, J. N. *Acc. chem. Res.* **20**, 415–421 (1987).
- Ribarsky, M. W. & Landman, U. *J. chem. Phys.* **97**, 1937–1949 (1992).
- Luedtke, W. D. & Landman, U. *Comput. Mater. Sci.* **1**, 1–24 (1992).
- Landman, U., Luedtke, W. D., Ouyang, J. & Xia, T. K. *Jap. J. appl. Phys.* **32**, 1444–1462 (1993).
- Rhykerd, C., Schoen, M., Diester, D. & Cushman, J. *Nature* **330**, 461–463 (1987).
- Schoen, M., Diester, D. & Cushman, J. *J. chem. Phys.* **87**, 5464 (1987); *Phys. Rev.* **B47**, 5603–5613 (1987).
- Diester, D., Schoen, M. & Cushman, J. *Science* **262**, 545–547 (1993).
- Han, K. K., Cushman, J. H. & Diester, D. *J. Molec. Phys.* **79**, 537–545 (1993).
- Schoen, M., Rhykerd, C., Diester, D. & Cushman, J. *Science* **245**, 1223–1225 (1989).
- Callaway, M., Tildesley, D. J. & Quirke, N. *Langmuir* **10**, 3350–3356 (1994).
- Wang, Y., Hill, K. & Harris, J. G. *J. chem. Phys.* **100**, 3276–3285 (1994).
- Kirk Hill, Y. W. & Harris, J. G. *Langmuir* **9**, 1983–1985 (1993).
- Gupta, S., Koopman, D. C., Westerman-Clark, G. B. & Bitsanis, I. A. *J. chem. Phys.* **100**, 8444–8453 (1994).
- Thompson, P. A. & Robbins, M. O. *Science* **250**, 792–794 (1990).
- Israelachvili, J. N. et al. *J. Phys.: Condensed Matter* **2**, SA89–SA98 (1990).
- Yoshizawa, H. & Israelachvili, J. *J. phys. Chem.* **97**, 11300–11313 (1993).
- Van Alsten, J. & Granick, S. *Phys. Rev. Lett.* **61**, 2570–2573 (1988).
- Hu, H. W., Carson, G. & Granick, S. *Phys. Rev. Lett.* **66**, 2758–2761 (1991).
- Reiter, G., Demirel, A. L. & Granick, S. *Science* **263**, 1741–1744 (1994).
- Xia, T. K., Ouyang, J., Ribarsky, M. W. & Landman, U. *Phys. Rev. Lett.* **69**, 1967–1970 (1992).
- Xia, T. K. & Landman, U. *Phys. Rev.* **B48**, 11313–11314 (1993).
- Xia, T. K. & Landman, U. *J. chem. Phys.* **101**, 2498–2507 (1994).
- Xia, T. K. & Landman, U. *Science* **261**, 1310–1312 (1993).
- Landman, U., Luedtke, W. D. & Ringer, E. M. *Wear* **153**, 3–30 (1992).
- Homola, A. M., Israelachvili, J. N., Gee, M. L. & McGuiggan, P. M. *ASME J. Tribology* **111**, 675–682 (1989).
- Robbins, M. O. & Thompson, P. A. *Science* **253**, 916 (1991).
- Gee, M. L., McGuiggan, P. M., Israelachvili, J. N. & Homola, A. J. *chem. Phys.* **93**, 1895–1906 (1990).
- Persson, B. N. J. *Phys. Rev.* **B50**, 4771–4786 (1994).
- Yoshizawa, H., Chen, Y. L. & Israelachvili, J. *J. phys. Chem.* **97**, 4128–4140 (1993); *J. Adhesion Sci. Technol.* **8**, 1–18 (1994).
- Yoshizawa, H., McGuiggan, P. & Israelachvili, J. N. *Science* **259**, 1305–1308 (1993).
- Klein, J., Kumacheva, E., Mahalu, D., Perahia, D. & Fetters, L. J. *Nature* **370**, 634–636 (1994).
- Blackman, G. S., Mate, C. M. & Philpott, M. R. *Phys. Rev. Lett.* **65**, 2270–2273 (1990).
- Burnham, N. A., Dominguez, D. D., Mowery, R. L. & Colton, R. J. *Phys. Rev. Lett.* **64**, 1931–1934 (1990).
- Landman, U., Luedtke, W. D., Burnham, N. A. & Colton, R. J. *Science* **248**, 454–461 (1990).
- Ducker, W. A., Senden, T. J. & Pashley, R. M. *Langmuir* **8**, 1831–1836 (1992).
- Salmeron, M. et al. *Langmuir* **9**, 3600–3611 (1993).
- Kaneko, R., Miyamoto, T. & Hamada, E. *Adv. Info. Storage Syst.* **1**, 267–277 (1991).
- Meyer, E. et al. *Thin Solid Films* **220**, 132–137 (1992).
- Bhushan, B. & Ruan, J. *ASME J. Tribology* **116**, 389–396 (1994).
- Bhushan, B., Koinkar, V. N. & Ruan, J. *Proc. Instn. Mech. Engrs.* **J208**, 17–29 (1994).
- Bhushan, B. & Koinkar, V. N. *J. appl. Phys.* **75**, 5741–5746 (1994).
- Ruan, J. & Bhushan, B. *J. appl. Phys.* **76**, 5022–5035 (1994).
- Ruan, J. & Bhushan, B. *J. appl. Phys.* **76**, 8117–8120 (1994).
- Bhushan, B. & Koinkar, V. N. *Tribology Trans.* **38**, 119–127 (1995).
- Bhushan, B. & Koinkar, V. N. *Wear* **180**, 9–16 (1995); **183**, 360–370 (1995).
- Miyamoto, T., Kaneko, R. & Miyake, S. *J. Vac. Sci. Technol.* **B9**, 1336–1339 (1991).
- Mate, C. M. *Phys. Rev. Lett.* **68**, 3323–3326 (1992).
- O'Shea, S. J., Welland, M. E. & Rayment, T. *Appl. Phys. Lett.* **61**, 2240–2242 (1992).
- Bhushan, B. et al. *Langmuir* (in the press).
- Bhushan, B., Miyamoto, T. & Koinkar, V. N. *Adv. Info. Storage Syst.* **6**, 151–161 (1995).
- Majumdar, A. & Bhushan, B. *ASME J. Tribology* **112**, 205–216 (1990).
- Majumdar, A. & Bhushan, B. *ASME J. Tribology* **113**, 1–11 (1991).
- Bhushan, B. & Majumdar, A. *Wear* **153**, 53–64 (1992).
- Ganti, S. & Bhushan, B. *Wear* **180**, 17–34 (1995).
- Burnham, N. A. & Colton, R. J. *J. Vac. Sci. Technol.* **A7**, 2906–2913 (1989).
- Maivald, P. et al. *Nanotechnology* **2**, 103–106 (1991).
- Bhushan, B. & Koinkar, V. N. *Appl. Phys. Lett.* **64**, 1653–1655 (1994).
- Landman, U., Luedtke, W. D. & Nitzan, A. *Surf. Sci.* **210**, L177–L184 (1989).
- Landman, U., Luedtke, W. D. & Ribarsky, M. W. *J. Vac. Sci. Technol.* **A7**, 2829–2839 (1989).
- Frisbie, C. D., Rozsnyai, L. F., Noy, A., Wrighton, M. S. & Lieber, C. M. *Science* **265**, 2071–2074 (1994).
- Gane, N. & Cox, J. M. *Phil Mag.* **22**, 881–891 (1970).
- Ruan, J. & Bhushan, B. *J. Mater. Res.* **8**, 3019–3022 (1993).
- Mate, C. M., Lorenz, M. R. & Novotny, V. J. *J. chem. Phys.* **90**, 7550–7555 (1989).
- Blackman, G. S., Mate, C. M. & Philpott, M. R. *Phys. Rev. Lett.* **65**, 2270–2273 (1990).
- Bhushan, B. & Blackman, G. S. *ASME J. Tribology* **113**, 452–458 (1991).
- Landman, U., Barnett, R. N., Cheng, H.-P., Cleveland, C. L. & Luedtke, W. D. in *Computations for the Nano-Scale 75–113* (eds Blochl, P. E., Joachim, C. & Fisher, A. J.) (Kluwer, Dordrecht, 1993).
- Landman, U., Luedtke, W. D. & Ringer, E. M. in *Fundamentals of Friction: Macroscopic and Microscopic Processes* 463–508 (eds Singer, J. L. & Pollock, H. M.) (Kluwer, Dordrecht, 1991).
- Landman, U. & Luedtke, W. D. *J. Vac. Sci. Technol.* **9**, 414–423 (1991).
- Thompson, P. A., Robbins, M. O. & Grest, G. S. in *Computations for the Nano-Scale 127–138* (eds Blochl, P. E., Joachim, C. & Fisher, A. J.) (Kluwer, Dordrecht, 1993).
- Luedtke, W. D. & Landman, U. *Phys. Rev. Lett.* **73**, 569–572 (1994).
- Dowson, D. *History of Tribology* (Longman, London, 1979).
- Pascual, J. I. et al. *Phys. Rev. Lett.* **71**, 1852–1855 (1993).
- Pascual, J. I. et al. *Science* **267**, 1793–1795 (1995); *J. Vac. Sci. Technol.* (in the press).
- Abraham, F. F., Rudge, W. E. & Alexopoulos, P. S. *Comput. mater. Sci.* **3**, 21–40 (1994).
- Harrison, J. A., White, C. T., Colton, R. J. & Brenner, D. W. *Mater. Res. Soc. Bull.* **17**, 50–53 (1993).
- Tupper, K. J., Colton, R. J. & Brenner, D. W. *Langmuir* **10**, 2041–2043 (1994).
- Tupper, K. J. & Brenner, D. W. *Langmuir* **10**, 2335–2338 (1994).
- Karaborni, S. *Phys. Rev. Lett.* **73**, 1668–1671 (1994).
- Thompson, P. A., Grest, G. S. & Robbins, M. O. *Phys. Rev. Lett.* **68**, 3448–3451 (1992).
- Landman, U. & Luedtke, W. D. *Mater. Res. Soc. Bull.* **17**, 36–44 (1993).
- Gao, J., Luedtke, W. D. & Landman, U. *Science* (submitted).
- Fowles, P. E. *J. Lub. Technol., Trans. ASME* **91**, 464–476 (1969).
- Li, X., Cheng, H. S. & Zheng, L., *ASME J. Tribology* **115**, 102–110 (1993).
- Chang, L., Jackson, A. & Webster, M. N. *Tribology Trans.* **37**, 435–444 (1994).

Glycogen synthase kinase-3 and dorsoventral patterning in *Xenopus* embryos

Xi He^{*}, Jean-Pierre Saint-Jeannet^{†‡}, James R. Woodgett[§],
Harold E. Varmus^{*} & Igor B. Dawid[†]

^{*} Varmus Laboratory, Bldg 49, Room 4A56, National Cancer Institute, [†] Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

[§] Ontario Cancer Institute, Toronto, Ontario M4X 1K9, Canada

Glycogen synthase kinase 3 (GSK-3) is homologous to the product of the *Drosophila* gene *shaggy* (*zeste-white 3*), which is required for signalling by *wingless* during *Drosophila* development. To test whether GSK-3 is also involved in vertebrate pattern formation, its role was investigated during early *Xenopus* development. It was found that dominant-negative GSK-3 mutants induced dorsal differentiation, whereas wild-type GSK-3 induced ventralization. These results indicate that GSK-3 is required for ventral differentiation, and suggest that dorsal differentiation may involve the suppression of GSK-3 activity by a *wingless*/wnt-related signal.

PATTERNING of the vertebrate embryo, as studied extensively in *Xenopus laevis*, depends on cell interactions during mesoderm induction and axis specification^{1,2}. The dorsoventral axis is first determined by a rotation of the egg cortex relative to the cytoplasm, which leads to the establishment of a dorsalizing or Nieuwkoop centre³. During subsequent cleavage, cell interactions specify the equatorial zone as future mesoderm and impart dorsoventral polarity upon it^{1,3}. At early gastrula, dorsal mesoderm (the Spemann organizer) has acquired the ability to establish the larval body plan through initiation and organization of gastrulation movements, formation of the notochord, neural induction and patterning, and induction of other dorsal mesodermal derivatives^{1,2}.

Major progress has been made in identifying signalling molecules in *Xenopus* embryogenesis^{1,2}. Mesoderm inducers include most members of the fibroblast growth factor (FGF) family, and members of the transforming growth factor- β (TGF- β) family, such as activin and BVgl, a processing-enhanced form of Vgl. Introduction of dominant-negative forms of FGF or activin receptor into developing embryos perturbs mesoderm formation, providing direct evidence that FGF, activin, or related molecules are required for mesoderm induction^{4,5}. Activin and BVgl also have properties of dorsalizing signals, and can partially (activin) or fully (BVgl) rescue dorsal axial structures in ventralized embryos generated by ultraviolet irradiation^{6,7}.

Certain members of the Wnt family of secreted glycoproteins⁸, notably mouse Wnt-1 (ref. 9) and *Xenopus* wnt-8 (Xwnt-8) (refs 10–12), and the unrelated factors noggin¹³ and chordin¹⁴, have characteristics of dorsalizing signals without being mesoderm inducers^{2,15}. Xwnt-8, either by eliciting generation of a Nieuwkoop centre or by mimicking a Nieuwkoop signal(s), can fully restore dorsal axis formation in ultraviolet-ventralized embryos, and can cause axis duplications when introduced into the ventral side of normal embryos^{10,11}. However, the expression patterns or the activities of the identified *Xenopus* wnt genes and that of noggin do not satisfy fully the expected requirements of the endogenous Nieuwkoop dorsalizing signal(s). Furthermore, in spite of the identification of potential dorsalizing factors, very little is known about the molecular nature of the signal transduc-

tion pathways that mediate natural or experimental dorsalizing influences.

Genetic studies in *Drosophila* have led to the identification of several molecules involved in the signalling cascade of wingless (wg), the homologue of Wnt-1. Among them is shaggy (sgg, also known as zeste-white 3, zw-3; ref. 16), a serine/threonine protein kinase homologous to mammalian glycogen synthase kinase-3 (GSK-3). GSK-3 was discovered as a kinase activity phosphorylating glycogen synthase and mediating insulin regulation of glycogen synthesis (reviewed in ref. 17), and has two isoforms, α and β , that are encoded by different genes¹⁸. GSK-3 is expressed ubiquitously¹⁸ and is believed to be a constitutively active kinase, whose activity is downregulated upon hormonal stimulation^{17,19,20}. Consistent with the biochemical studies of GSK-3, genetic evidence suggests that sgg/zw3 behaves as a constitutively active kinase and is inhibited by wg signalling²¹. Mammalian GSK-3 β can rescue the sgg/zw-3 mutant phenotype connected with wg signal transduction²¹ and with Notch signalling²². As wnt signals can induce dorsalization in *Xenopus*, we hypothesized that downregulation of GSK-3 activity may participate in axis formation.

Here we report studies on the function of GSK-3 in frog embryogenesis. Ventral injection of dominant-negative GSK-3 mutants induces dorsal development and duplication of the body axis, whereas injection of wild-type GSK-3 on the dorsal side ventralizes the embryo. Mutant GSK-3 fully rescues dorsal axis formation in ultraviolet-ventralized embryos by restoring a Nieuwkoop centre, whereas overexpression of wild-type GSK-3 blocks the dorsalizing ability of injected Xwnt-8 messenger RNA. These results suggest that GSK-3 regulates the establishment of dorsoventral polarity.

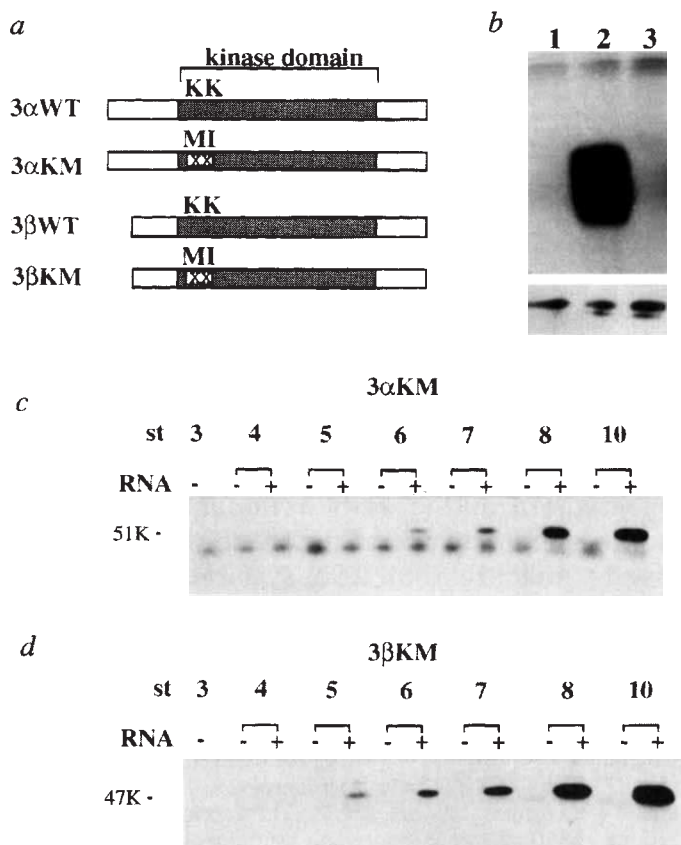
Dorsal axis induction by mutant GSK-3

Complementary DNAs encoding human GSK-3 α and GSK-3 β , referred to as 3 α WT and 3 β WT, and their corresponding kinase mutants, GSK-3 α KM (3 α KM) and GSK-3 β KM (3 β KM), are schematically represented in Fig. 1a. Both mutants were generated by altering the two consecutive lysine residues in the ATP-binding site to a methionine and an isoleucine. 3 β WT and 3 β KM proteins synthesized in mRNA-injected embryos were assayed *in vitro* for their kinase activities. As expected, 3 β WT was highly active in phosphorylating a GSK-3-specific substrate, whereas 3 β KM had no kinase activity (Fig. 1b). Studies with other protein kinases have shown that analogous ATP-binding site mutants are inactive and behave as inhibitors of the corre-

[‡] Permanent address: Centre de Biologie du Développement, UMR-CNRS 9925, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex, France.

FIG. 1. Schematic representation of GSK-3 constructs, *in vitro* kinase assay and protein accumulation in GSK-3KM RNA-injected embryos. **a**, Wild-type GSK-3 α (3 α WT) and GSK-3 β (3 β WT) and kinase mutants GSK-3 α (3 α KM) and GSK-3 β (3 β KM). Shaded boxes represent the kinase domains in which two consecutive lysine residues (KK) in the ATP-binding site were mutated into methionine and isoleucine residues (MI). These double mutations were designed to avoid any potential compensatory effects. **b**, Top, kinase activity of control immunoprecipitates (lane 1), and immunoprecipitated 3 β WT (lane 2) and 3 β KM (lane 3), using extracts prepared from stage-11 embryos injected with the corresponding mRNAs at the 4-cell stage. Bottom, immunoblot of the same samples. Upper band is mouse IgG heavy chain present in all lanes; lower band in lanes 2 and 3 corresponds to GSK-3 β , indicating approximately equal input of 3 β WT and 3 β KM proteins in the assay. **c** and **d**, Immunoblots of 3 α KM (**c**) and 3 β KM (**d**) proteins in embryo extracts prepared from different stages of uninjected (–) and injected (+) embryos. 3 α KM and 3 β KM RNAs were injected ventrally at the 4-cell stage (st. 3) at 10 ng per embryo. The relative molecular mass of GSK-3 α and GSK-3 β is 51K and 47K, respectively. On the autoradiogram, the protein levels of 3 β KM appeared to be five to ten times those of 3 α KM when equal exposure time was applied. Similar accumulation patterns were obtained with embryos injected with 3 α WT and 3 β WT RNAs.

METHODS. All GSK-3 constructs used for synthesizing RNA *in vitro* are based on pSP64T (ref. 45). Human GSK-3 α and GSK-3 β (accession number L40027 and L33801, respectively) in BlueScript vector (Stratagene) were used for site-directed mutagenesis according to ref. 46. *Bsp*HI and *Nco*I restriction sites were introduced at the initiation ATG codons of GSK-3 α and GSK-3 β , respectively. The creation of *Nco*I site in GSK-3 β changed the second amino acid from serine to alanine. *Bst*EII restriction sites were introduced 3' to the stop codons. 3 α KM was generated by changing lysines 128 and 129 in 3 α WT into methionine and isoleucine; 3 β KM was generated by similarly altering lysines 85 and 86. Wild-type and mutant coding regions were subcloned into pSP64T between the *Nco*I and *Bst*EII sites, and capped RNAs were transcribed as described⁴⁵. For *in vitro* kinase assay, we used constructs tagged at the C termini with the haemagglutinin epitope (HA). 0.5 ng RNA for 3 β WT or 3 β KM per blastomere (1 ng per embryo) was injected near the ventral equatorial zone at the 4-cell stage. Extracts were made by homogenizing 10 embryos at stage 11 in 400 μ l cold NP-40 buffer (137 mM NaCl, 50 mM NaF, 0.5% NP-40, 5 mM EDTA, 10 mM Tris, pH 7.5, 1 mM NaVO₄, 2 mM PMSF, 25 mM leupeptin, 0.2 U ml^{–1} aprotinin)²⁶, and were immunoprecipitated with 10 μ g monoclonal antibody (12CA5) against HA epitope and 20 μ l of 50% protein A–Sepharose beads (Sigma). Half of the precipitates was used for kinase assay with a GSK-3-specific substrate as described in ref. 20, except that the phosphopeptide was resolved on a 25% SDS–PAGE. The other half was used for immunoblotting with a specific antisera against GSK-3 β (see below). For protein accumulation



analysis, 5 ng RNA per blastomere (10 ng per embryo) was injected and embryonic extracts were made as described. 30 μ l of each extract (0.7 embryo equivalent) at different stages were separated on 12% SDS–PAGE, transferred to a 0.2 μ m nitrocellulose filter (Schleicher & Schuell) and blocked with 2% BSA in TBST (10 mM Tris, pH 7.5, 150 mM NaCl, 0.2% Tween 20). Antisera specific for human GSK-3 α or 3 β were used at 1:1,000 in 1% BSA in TBST, incubated overnight at 4 $^{\circ}$ C, and visualized using peroxidase-conjugated secondary antibody (Boehringer–Mannheim) and ECL detection system (Amersham). The intensity of reactive bands in immunoblots reflected protein levels, because identical intensities resulted when equal amounts of recombinant GSK-3 α and GSK-3 β were blotted by their respective antisera (not shown).

sponding wild-type kinases, presumably by blocking access to substrates and/or regulatory factors (see refs 23–26 for example). Therefore, we expected that 3 α KM and 3 β KM proteins would act as dominant inhibitors of endogenous GSK-3.

The accumulation of 3 α KM and 3 β KM protein products after mRNA injection was examined by immunoblotting at different stages of development. 3 α KM (Fig. 1c) and 3 β KM (Fig. 1d) proteins were detected within 30 min of injection (8-cell, stage 4), and continued to accumulate at least up to stage 10 (early gastrula). For the same amount of injected RNA, 3 β KM protein accumulated to levels 5- to 10-fold those of 3 α KM (Fig. 1c, d).

Ventral injection at the 4-cell stage of either 3 α KM or 3 β KM RNA induced the formation of ectopic dorsal axes, detectable as early as stage 19 (Fig. 2a). The induced axes were frequently complete, with well-defined eyes and cement glands, such that original and induced axes were hardly distinguishable (Fig. 2b). 3 β KM was more potent in its ability to induce supernumerary axes; embryos with triplicated complete axes were observed, with excessive head development (including three cement glands and four to six eyes) at the expense of posterior structures (Fig. 2a, b, d). Histological examination of embryos injected with 3 β KM RNA demonstrated axis triplication, including three

notochords surrounded by somites and three neural tubes (Fig. 2c). In the case of 3 α KM, the incidence and completeness of axis duplication were dose dependent (Fig. 2d). While 5–10 ng of 3 α KM RNA was required for complete axis duplication, 1 ng of 3 β KM RNA was sufficient, probably due to the difference in protein levels (Fig. 1c, d). Ectopic formation of dorsal axes was strictly dependent on the site of injection; embryos injected dorsally with 3 α KM or 3 β KM RNAs (10 ng per embryo) developed normally (Fig. 2b). Wild-type GSK-3 injected ventrally did not induce axis duplication (see below).

To test the specificity of the effects of mutant GSK-3, we co-injected RNAs for 3 α KM or 3 β KM with RNAs for 3 α WT or 3 β WT into the ventral equatorial region of 4-cell embryos. Co-injection of 3 α WT with 3 α KM or 3 β KM RNAs largely blocked the induction of secondary body axes and rendered the majority of injected embryos indistinguishable from uninjected controls (Fig. 3a, b). Likewise, 3 β WT RNA was able to block the formation of secondary axes induced by 3 β KM or by 3 α KM (Fig. 3b). These results support the view that mutant GSK-3 affects embryonic development by specifically interfering with endogenous GSK-3 activity. In addition, co-injection of RNA for the ventral inducer XBMP-4 (*Xenopus* bone morphogenetic protein-4; refs 27, 28) with the mutant GSK-3 on the ventral side blocked

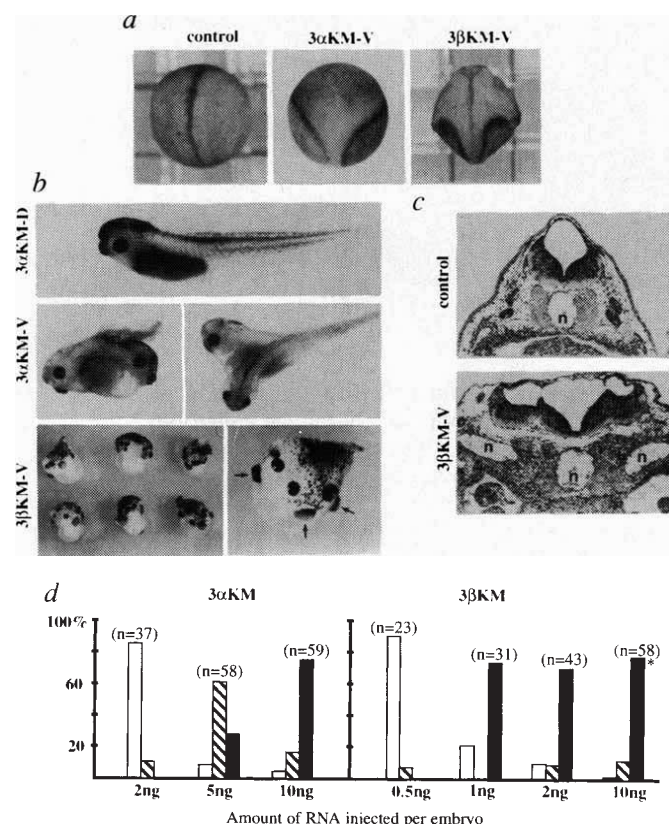


FIG. 2 Dorsal axis duplication induced by kinase mutants 3α KM and 3β KM. **a**, Uninjected embryo (left), and embryos injected ventrally at the 4-cell stage with 10 ng (5 ng per blastomere) of 3α KM (middle) and 3β KM RNA (right), shown at the neurula stage (stage 19). Double and triple axis formation were evident in the 3α KM- and the 3β KM-injected embryos, respectively. **b**, Embryos (all shown at stage 41) injected dorsally with 3α KM RNA (10 ng per embryo) were normal (top). 3α KM-ventrally injected embryos (middle) show complete axis duplication including eyes and cement gland. 3β KM-ventrally injected embryos (bottom) display up to three cement glands (arrows) and five eyes. **c**, Transverse sections of a control embryo (top) and an embryo injected ventrally with 3β KM (10 ng per embryo, bottom). Note the presence in 3β KM-injected embryo of three neural tubes and notochords (n), and somites surrounding each notochord. **d**, Histogram of frequency of axis duplication in embryos injected ventrally with different concentrations of 3α KM and 3β KM RNAs. n, Total number of embryos examined. White bar, normal embryo; grey bar, incomplete axis duplication (lacking the anterior-most structures); hatched bar, complete axis duplication. Asterisk, for injection of 10 ng of 3β KM RNA, embryos presenting triplication (28%) or duplication (52%) of the axis have been combined. **METHODS**. Eggs were obtained and embryos were raised in $0.1 \times$ normal amphibian medium (NAM) and staged according to ref. 47. RNA injections were done in 5% Ficoll, $0.1 \times$ NAM, at the 4-cell stage, with two blastomeres injected with equal amounts of RNA. Unless otherwise specified, injections were at the equatorial zone near the midline separating the two dorsal or ventral blastomeres. Dorsal blastomeres were recognized by their lighter pigmentation (100% accuracy was achieved using β -galactosidase as lineage tracer). Fixation, sectioning and staining have been described¹².

secondary dorsal axis induction (Fig. 3c), indicating that the XBMP-4-dependent ventralizing pathway could function in the presence of GSK-3KM. This also suggests that XBMP-4 acts downstream of GSK-3 in dorsoventral patterning.

Ventralization by wild-type GSK-3

The fact that 3α KM and 3β KM RNAs did not affect development when injected dorsally suggests that GSK-3 is either absent or its activity is already inhibited on the dorsal side of the

embryo. Overexpression of wild-type GSK-3 in the dorsal blastomeres thus may cause ventralization. When 3α WT or 3β WT RNAs were injected into the equatorial region of the two dorsal blastomeres at the 4-cell stage, the resulting embryos were significantly ventralized. These embryos, in many cases, lacked anterior structures including eyes and forebrain, but always retained a posterior dorsal axis, namely trunk and tail (Fig. 3d). Ventralization was dose dependent for both kinases as expressed in the dorsoanterior index, DAI (Fig. 3e, f); it should be noted that the DAI is not a linear or quantitative scale, but a means for scoring levels of development²⁹. For the same amount of RNA injected, 3β WT appeared to be more effective in its ventra-

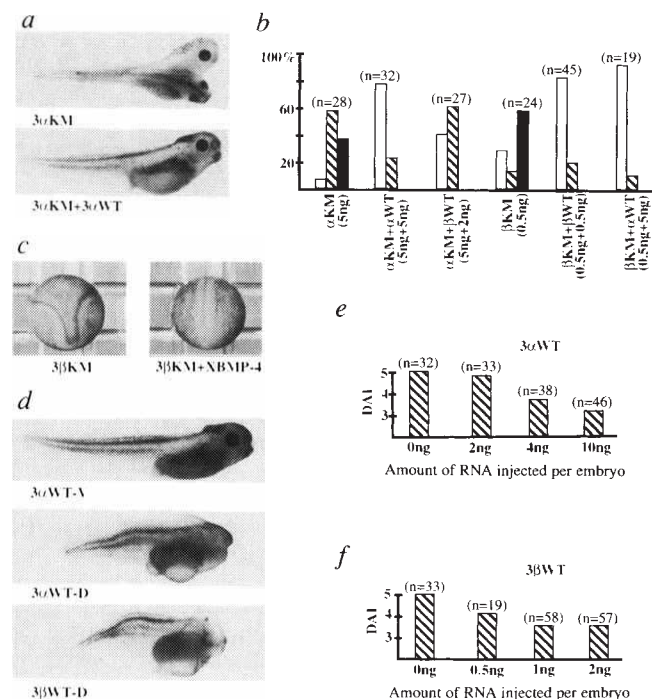
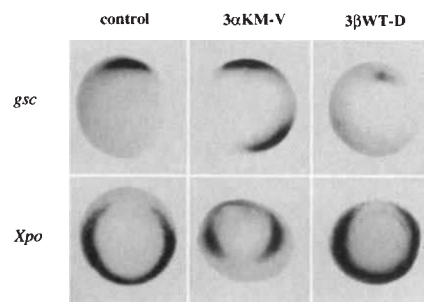


FIG. 3 Wild-type GSK-3 (and XBMP-4) suppresses dorsalization by mutant GSK-3 and ventralizes normal embryos. **a**, Co-injection of mRNAs of wild-type and mutant kinases. Top, embryo injected ventrally with 3α KM RNA (5 ng per embryo); bottom, embryo co-injected with 3α KM and 3α WT RNAs, each at 5 ng per embryo. **b**, Histogram of axis duplication frequency obtained by co-injection of wild-type and mutant GSK-3 RNA. Note that axis duplication was suppressed in all co-injection experiments. n, Total number of embryos examined. White bar, normal embryo; grey bar, incomplete axis duplication; hatched bar, complete axis duplication. **c**, Embryos injected ventrally with 3β KM RNA (2 ng per embryo) or with 3β KM RNA (2 ng per embryo) plus XBMP-4 RNA (400 pg per embryo). When scored at stage 19 (shown) or at stage 41, 98% ($n=27$) of the 3β KM RNA-injected embryos exhibited double axes, but only 8% ($n=49$) of the embryos co-injected with 3β KM and XBMP-4 RNAs had axis duplication. Co-injection with less XBMP-4 RNA (100 pg per embryo) inhibited axis duplication as well (30% axis duplication, $n=29$). As a control for XBMP-4 activity, dorsal injection of the same concentration of XBMP-4 RNA resulted in ventralized embryos, as reported^{27,28}. **d**, Embryos injected ventrally (top) or dorsally (centre) with 3α WT RNA (4 ng per embryo), or dorsally with 3β WT RNA (1 ng per embryo) (bottom). Ventrally injected embryos were normal, but dorsally injected embryos showed anterior truncation of the body axis (DAI, 2). **e** and **f**, Average dorsoanterior index (DAI) of embryos injected dorsally with different concentrations of 3α WT RNA (**e**) or 3β WT RNA (**f**). In the DAI²⁹, a value of 5 represents normal embryos, and DAI of 0 signifies complete ventralization with no axis. Embryos with DAI of 3 have severe loss of anterior features such as forebrain, eyes and cement gland. Injection of 4 ng or more 3β WT RNA per embryo resulted in gastrulation defects and lethality. **METHODS**. XBMP-4 cDNA⁴⁸ was subcloned into the SP64T vector, and RNA was transcribed as previously described⁴⁵.

FIG. 4 Expression of GSK-3WT and GSK-3KM affects marker gene expression in the early gastrula, as seen by whole-mount *in situ* hybridization of *goosecoid* (*gsc*) and *Xpo*. Control, uninjected embryos; 3 α KM-V, embryos injected ventrally with 3 α KM RNA (10 ng per embryo); 3 β WT-D, embryos injected dorsally with 3 β WT RNA (1 ng per embryo). The organizer-specific homeobox gene *gsc*³⁰, and *Xpo*, a zinc-finger gene specific for ventrolateral mesoderm³¹, were used as probes. Vegetal views are shown for all embryos, with the original dorsal side pointing up. Ventrally injected 3 β KM RNA induced similar or higher levels of ectopic *gsc* expression. Dorsally injected 3 α WT RNA showed similar results to 3 β WT, though higher amounts of RNA were required. METHODS. Antisense probes were synthesized *in vitro* using digoxigenin-UTP. Whole-mount *in situ* hybridization was done as described⁴⁹.



lizing ability than 3 α WT (Fig. 3*e,f*), probably as a result of higher amounts of protein produced. Ventral injection of RNAs for 3 α WT (up to 10 ng per embryo; Fig. 3*d*, top) and 3 β WT (up to 2 ng per embryo; not shown) had no obvious effect on embryogenesis.

Spemann organizer formation

At the early gastrula stage the dorsoventral polarity of the embryo becomes overt with the formation of the Spemann organizer. To investigate early effects of the expression of GSK-3 and its mutants on the organization of the body plan, we carried out whole-mount *in situ* hybridization on gastrula embryos with two markers: *goosecoid* (*gsc*), a homeobox gene that specifically marks the dorsal mesoderm, and *Xpo*, a zinc-finger gene with restricted expression in the ventrolateral mesoderm^{30,31} (Fig. 4, control). In embryos injected ventrally with 3 α KM RNA, ectopic induction of *gsc* and local inhibition of *Xpo* were seen opposite the original dorsal side of the embryo (Fig. 4, centre). In embryos injected dorsally with 3 β WT RNA, endogenous *gsc* expression was greatly suppressed, whereas *Xpo* expression expanded dorsally (Fig. 4, right). These results suggest that ventral injection of the mutant GSK-3 induced an ectopic organizer in the area that would have developed as ventral mesoderm, whereas dorsal injection of wild-type GSK-3 interfered with the formation of the endogenous organizer, fully consistent with the phenotypes seen at later stages of development.

Rescue of UV-ventralized embryos

Ultraviolet irradiation of embryos during the first cell cycle blocks cortical rotation and subsequent formation of the Nieuwkoop centre, resulting in ventralized embryos³. Artificially generated rotation or transplantation of dorsal vegetal cells into ultraviolet-treated embryos can fully restore dorsal development, suggesting that the irradiated embryos lack only the proper dorsalizing signals³. To test whether suppression of GSK-3 activity can rescue dorsal development in ultraviolet-ventralized embryos, 5 ng of 3 α KM or 1 ng of 3 β KM RNA was injected into the equatorial region of two blastomeres at the 4-cell stage. As shown in Fig. 5, either 3 α KM (Fig. 5*b,c*) or 3 β KM (Fig. 5*c,d*) RNA was able to restore axial development in treated embryos, with many of the rescued embryos being indistinguishable from controls.

Generation of a Nieuwkoop centre

We next investigated how axis rescue is achieved. Dorsal axis patterning involves multiple signalling steps, from the generation of the Nieuwkoop centre during early cleavage to the formation of the Spemann organizer in the later blastula¹⁻³. In an operational definition, the Nieuwkoop centre induces a dorsal axis without contributing progeny cells to axial structures, whereas the Spemann organizer both differentiates into axial structures and recruits other cells to participate. To determine how 3 α KM and 3 β KM function in these multistep induction events, we performed lineage tracing experiments by co-injection of RNAs for β -galactosidase (β -gal) and 3 α KM or 3 β KM into ultraviolet-

treated embryos. To direct the RNAs to different regions of the embryo, injections were aimed at the equatorial zone or near the vegetal pole (Fig. 5*a* and *d*, legend). Co-injection of β -gal RNA with 3 α KM or 3 β KM RNA near the equatorial zone always led to X-gal staining in the rescued axis and endodermal derivatives (Fig. 5*d*, upper panel). Injection near the vegetal pole was somewhat less effective in axial rescue, but a majority of injected embryos generated axial structures, some of which were complete, including anterior head structures (Fig. 5*d*, lower panel). In the rescued embryos, the lineage label was absent from axial tissue in 72% of the cases (Fig. 5*d*, lower panel), whereas the remaining embryos showed X-gal staining in both endoderm and axial structures. Therefore, 3 α KM- or 3 β KM-injected cells could promote axial rescue without generating tissues like notochord, somites, brain or spinal cord. These results suggest that cells expressing the mutant kinase were the source of axis-inducing signals, that is, they could behave as a Nieuwkoop centre.

Suppression of axis induction by Xwnt-8

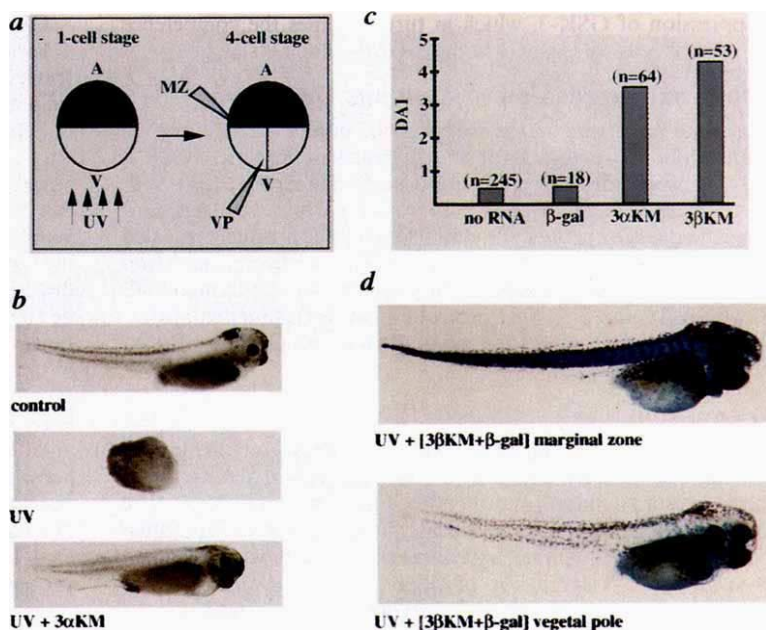
Injection of Xwnt-8 mRNA is capable of inducing a complete dorsal axis in *Xenopus* embryos^{10,11}. If, as we propose, the Xwnt-8 signal is transduced by inactivating GSK-3, overexpression of the wild-type kinase may override this inactivation and prevent dorsalization by Xwnt-8. We tested this prediction in two ways. The effective rescue of ultraviolet-ventralized embryos by injection of Xwnt-8 RNA was almost completely suppressed by co-injection of 1 ng 3 β WT RNA (Fig. 6*a,b*). Likewise, complete axis duplication induced by Xwnt-8 RNA was largely suppressed by coexpression of 1 ng 3 β WT RNA (Fig. 6*c*). These results are consistent with the view that GSK-3 is a component of the wnt signal transduction pathway in *Xenopus* embryos.

Sensitization of animal cap cells to FGF

A further similarity between the action of Xwnt-8 and 3 β KM in embryos is evident from their effects in animal cap cells. These cells are fated to become ectoderm but can be induced to form mesoderm. Ectopic expression of Xwnt-8 in animal explants does not induce mesoderm differentiation *per se*, but sensitizes the animal cap cells so that basic (b)FGF, normally a ventral mesodermal inducer, can now elicit the formation of dorsal mesodermal derivatives¹⁵. As shown in Fig. 6*d,e,f*, animal caps from both uninjected and 3 β KM RNA-injected embryos formed spherical structures characteristic of epidermal differentiation, and did not express skeletal muscle or neural markers we examined. Upon treatment with bFGF, uninjected animal caps responded with only minor deformation, and remained negative with these two markers. However, 3 β KM-injected caps elongated extensively when treated with bFGF, a morphological change associated with induction of dorsal mesoderm (Fig. 6*d*); similar elongation was observed in caps treated with the dorsal mesoderm inducer activin¹⁵. Upon immunostaining, these caps expressed the skeletal muscle and neural markers (Fig. 6*e,f*), indicative of dorsal type mesoderm formation and secondary neural induction. This result supports the view that Xwnt-8 signalling in the animal region, as in the equatorial zone, involves

FIG. 5 Rescue of UV-ventralized embryos and lineage tracing of injected cell progeny. **a**, Scheme of experimental design. A, animal pole; V, vegetal pole. Embryos were irradiated vegetally with UV before first cleavage, and were injected with 3 α KM or 3 β KM RNAs at the equatorial or marginal zone (MZ) at the 4-cell stage. For lineage tracing, β -galactosidase (β -gal) RNA was co-injected with 3 α KM or 3 β KM RNA either at the equatorial zone or near the vegetal pole (VP). **b**, 3 α KM rescue of UV-ventralized embryos. All embryos shown at equivalent stage 41. Control, normal embryo; UV, UV-ventralized embryo; UV + 3 α KM, UV-treated embryo rescued by 3 α KM (10 ng RNA per embryo), displaying a DAI between 4 and 5. **c**, Average DAI of UV-ventralized embryos and embryos rescued by equatorial injection of 3 α KM or 3 β KM RNA. Injection of β -galactosidase RNA had no effect. **n**, Total number of embryos examined. **d**, Lineage tracing. Embryos rescued by equatorially injected 3 β KM RNA ($n=46$) showed β -gal staining in axial structures and part of the endodermal derivatives in all cases (top). By contrast, in a majority (72%) of embryos rescued by vegetally injected 3 β KM RNA ($n=44$), β -gal staining was absent from axial structures (bottom); in the remaining 28% of cases, both axial and non-axial structures were stained. Vegetally injected embryos consistently showed less X-gal staining as compared to those injected at the equatorial zone, possibly reflecting a lower translation efficiency of the injected RNA. Control embryos injected with β -gal RNA in the dorsal equatorial zone ($n=18$) showed staining in the dorsal axis; when injected in dorsal vegetal region ($n=32$), 69% stained only in endoderm and the rest in endoderm and axis.

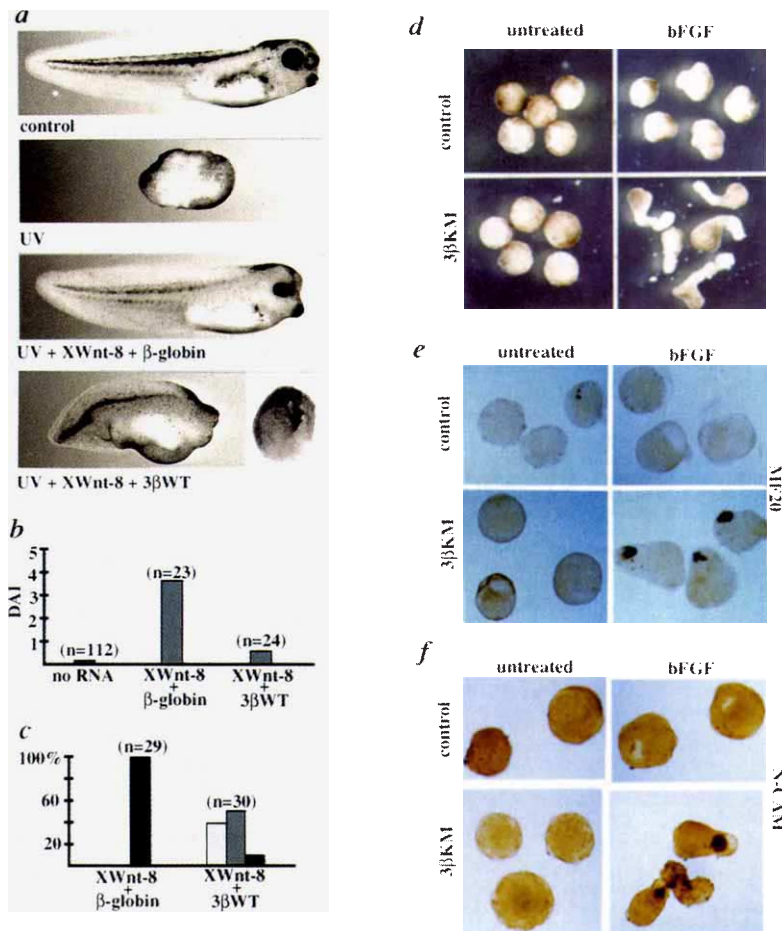
METHODS. Embryos were UV-irradiated within 30 min of fertilization in a UV crosslinker (Stratagene) with energy metre levels of 40–60 mJ.



For lineage tracing, 1 ng RNA for nuclear localized β -gal (transcribed from plasmid CS2 + n β -gal; ref. 50) was co-injected with 10 ng 3 α KM or 2 ng 3 β KM RNA. The distribution of β -gal was examined by X-gal staining¹⁰.

FIG. 6 3 β WT can suppress Xwnt-8 signalling in whole embryos, and 3 β KM mimics Xwnt-8 signalling in animal caps. **a**, Control, UV-treated embryo, and UV-treated embryos injected with 1 pg Xwnt-8 and 1 ng of β -globin RNAs, or with 1 pg Xwnt-8 and 1 ng 3 β WT RNAs; in the bottom panel, two examples of more or less ventralized embryos are shown. **b**, Average DAI of UV-treated embryos (DAI 0.1), UV-treated embryos injected with Xwnt-8 plus β -globin RNAs (DAI 3.4), or Xwnt-8 plus 3 β WT RNAs (DAI 0.6). **c**, Normal embryos were injected at the 4-cell stage into the ventral equatorial zone with 1 pg Xwnt-8 RNA in combination with β -globin or 3 β WT RNA (1 ng). Formation of secondary axes was scored as complete including eyes and cement gland (black bar), partial (grey bar) or no secondary axis (light bar). Although 100% of embryos injected with Xwnt-8 and β -globin RNAs formed complete secondary axes, only 10% of embryos co-injected with 3 β WT RNA did so. **d, e, f**, Animal caps excised from blastulae derived from uninjected embryos (top panels) or embryos injected with 5 ng 3 β KM RNA into the animal region (bottom panels). Note strong elongation (**d**, bottom right), illustrated at equivalent stage 18. Immunostaining with antibody MF20 for skeletal muscle myosin (**e**), and with antibody 4d against neural cell adhesion molecule N-CAM (**f**), at equivalent stage 25, were positive in 60% of 3 β KM-injected caps treated with bFGF but negative in all control caps. Staining with the notochord marker Tor70 was negative in all the explants examined. Stage-25 embryos were used as control for all immunostaining.

METHODS. Animal caps were excised from blastulae (stage 8–9) of control embryos (uninjected) or embryos injected at the 2-cell stage with 5 ng 3 β KM RNA into the animal region. Explants were cultured for 1 h in control medium (NAM, 0.5 $\times$) or in the presence of 100 ng ml⁻¹ bFGF (BRL), and were then transferred to 0.5 $\times$ NAM and further cultured to the equivalent stages already listed. Activin A-treated (50 pM) explants were included as positive controls (not shown). Whole-mount immunostaining was done as described¹⁵ using monoclonal antibodies MF20, 4d (both from Developmental Studies Hybridoma Bank, NICHD/NIH) and Tor70 (gift from R. Harland and P. Kushner).



suppression of GSK-3, which in turn modifies the competence of animal cells to respond to mesoderm-inducing signals.

Maternal expression of *Xenopus* GSK-3

As initial patterning in the frog embryo occurs without zygotic transcription, we tested for the presence of maternal GSK-3 RNAs by performing polymerase chain reaction (PCR) with degenerate primers, using cDNA from unfertilized eggs as template. Sequence analysis of the subcloned PCR products revealed the existence of at least two *Xenopus* GSK-3 isoforms (accession numbers U22890, U22891) that are 95–99% similar to mammalian GSK-3 α and β . RNase protection assays further demonstrated the maternal expression of both *Xenopus* GSK-3 mRNAs (data not shown).

Discussion

Dorsoventral patterning in the *Xenopus* embryo requires multiple inductive events, involving Nieuwkoop centre action, mesoderm induction and formation of the Spemann organizer^{1,3}. The experiments reported here implicate the serine/threonine protein kinase GSK-3 in dorsal axis formation. As blocking GSK-3 induces and overexpression of GSK-3 suppresses dorsal differentiation, inhibition of GSK-3 appears necessary and sufficient to initiate dorsal axis specification. The reciprocity of these results strongly argues for a role of GSK-3 in axial patterning. Preliminary results suggest equal distribution of the mRNAs for the two maternal GSK-3 isoforms along the presumptive dorsoventral axis (not shown). Although protein distribution is unknown, we speculate that the regulation of kinase activity, rather than protein concentration, is the main determinant. GSK-3 seems to control an early step in dorsalization, since cells in which GSK-3 was inhibited could function as a Nieuwkoop centre by inducing an axis without contributing cells to its formation.

The effect of the dominant-negative GSK-3 mutants on frog embryogenesis is similar to that of Xwnt-8 (refs 10, 11) or Wnt-1 (ref. 9), whereas wild-type GSK-3 counteracts the wnt effect. In parallel with the evidence that wg functions by inhibiting *sgg/zw3* in *Drosophila*²¹, these results suggest that the dorsalizing function of Xwnt-8 or Wnt-1 is mediated by inhibition of GSK-3. The wg signalling cascade in *Drosophila* includes gene products encoded by *dishevelled* and *armadillo*, and epistatic relationships suggest that they function upstream and downstream of *sgg/zw3*, respectively^{32,33}. A vertebrate Armadillo homologue, β -catenin, is implicated in patterning in *Xenopus* embryos, because injection of an antibody against it can cause axis duplication³⁴ and its depletion by antisense oligodeoxynucleotides suppresses dorsal development³⁵.

GSK-3 and *sgg/zw3* are implicated in signal transduction by a variety of stimuli in addition to wg/wnt, including insulin (reviewed in ref. 17) and Notch, a membrane protein mediating cellular communication²². In *Xenopus* embryos, several families of peptide growth factors other than wnts, such as noggin, activin, BVgl and chordin can function as dorsalizing signals^{1,2,7,14,36}. The action of these factors differs in that Xwnt-8, noggin and chordin do not induce mesoderm, but dorsalize ventral mesoderm^{1,2,14,15} and sensitize cells to mesoderm-inducing factors¹⁵, whereas BVgl and activin induce dorsal mesoderm directly^{1,2,7,36}. Two interpretations, not mutually exclusive, may be suggested. (1) There may exist two separate signal transduction pathways, one in mesoderm induction involving activation of ras-raf function^{26,37}, the other in dorsoventral patterning involving inhibition of GSK-3. Activin and BVgl may stimulate both pathways and thus function as both mesoderm inducers and dorsalizing signals, whereas wnts, noggin and chordin might activate only the second pathway. (2) As cells in which GSK-3 is inhibited can function as a Nieuwkoop centre, wnts might mimic signals inducing the Nieuwkoop centre rather than signals emanating from it. Inhibition of GSK-3 would stimulate the production of such Nieuwkoop signals, among which Vgl remains a candidate^{7,36}. In *Drosophila*, wg function in axial patterning often requires the synthesis of *decapentaplegic* (for example, see ref. 38), which, like Vgl, is a member of the TGF- β family. Dorsoventral patterning also relies upon active ventralizing signals, as suggested by the function of BMP-4, also a member of the TGF- β family^{27,28,39,40}. As BMP-4 blocks the dorsalizing action of the dominant-negative GSK-3 mutant, BMP-4 appears to function downstream of GSK-3 in dorsoventral axis specification.

A striking similarity emerges from a comparison of dorsoventral axis formation in the *Xenopus* embryo and the insect leg. The ventral region of the leg acts as an organizer, as seen in grafting experiments in cockroaches⁴¹. In the *Drosophila* imaginal disc, wg expression is restricted to the ventral region and ectopic expression of wg mimics organizer function and induces axis duplication^{38,42}. Importantly, clonal cells lacking *sgg/zw3* behave as organizers, inducing ventralization of dorsal cells and formation of secondary leg axes⁴³. Thus, the wg signal appears to promote the formation of an organizer by inhibiting *sgg/zw3* activity. The wg/wnt signalling cascade and the regulation of *sgg/zw3* and GSK-3 represent a conserved strategy during evolution for establishing embryonic polarity of both invertebrates and vertebrates. These studies, together with the finding that GSK-3 regulates cell fate in *Dictyostelium*⁴⁴, illustrate the essential developmental functions of the GSK-3 family in multicellular organisms. □

Received 27 December 1994; accepted 13 March 1995.

- Dawid, I. B. *J. biol. Chem.* **269**, 6259–6262 (1994).
- Kessler, D. S. & Melton, D. A. *Science* **266**, 596–604 (1994).
- Gerhart, J. C. *et al. Development (suppl.)* **107**, 37–51 (1989).
- Amaya, E., Musci, T. J. & Kirschner, M. W. *Cell* **66**, 257–270 (1991).
- Hemmati-Brivanlou, A. & Melton, D. A. *Nature* **359**, 609–614 (1992).
- Thomsen, G. *et al. Cell* **63**, 485–493 (1990).
- Thomsen, G. H. & Melton, D. A. *Cell* **74**, 433–441 (1993).
- Nusse, R. & Varmus, H. A. *Cell* **69**, 1073–1087 (1992).
- McMahon, A. P. & Moon, R. T. *Cell* **58**, 1075–1084 (1989).
- Smith, W. C. & Harland, R. M. *Cell* **67**, 753–765 (1991).
- Sokol, S., Christian, J. L., Moon, R. T. & Melton, D. A. *Cell* **67**, 741–752 (1991).
- Christian, J. L. & Moon, R. T. *Genes Dev.* **7**, 13–28 (1993).
- Smith, W. C. & Harland, R. M. *Cell* **70**, 829–840 (1992).
- Sasai, Y. *et al. Cell* **79**, 779–790 (1994).
- Christian, J. L., Olson, D. J. & Moon, R. T. *EMBO J.* **11**, 33–41 (1992).
- Perrimon, N. *Cell* **76**, 781–784 (1994).
- Woodgett, J. R. *Sem. Cancer Biol.* **5**, 269–275 (1994).
- Woodgett, J. R. *EMBO J.* **9**, 2431–2448 (1990).
- Sutherland, C., Leighton, I. A. & Cohen, P. *Biochem. J.* **296**, 15–19 (1993).
- Stambolic, V. & Woodgett, J. R. *Biochem. J.* **303**, 701–704 (1994).
- Siegfried, E., Chou, T. B. & Perrimon, N. *Cell* **71**, 1167–1179 (1992).
- Ruel, L., Bourouis, M., Heitzler, P., Pantescio, V. & Simpson, P. *Nature* **362**, 557–560 (1993).
- Han, M., Golden, A., Han, V. & Sternberg, P. W. *Nature* **363**, 133–140 (1993).
- Dumont, D. *et al. Genes Dev.* **8**, 1897–1909 (1994).
- Ogura, K. *et al. Genes Dev.* **8**, 2389–2400 (1994).
- MacNicol, A. M., Muslin, A. J. & Williams, L. T. *Cell* **73**, 571–583 (1993).
- Dale, L., Howes, G., Price, B. M. & Smith, J. C. *Development* **115**, 573–585 (1992).
- Jones, C. M., Lyons, K. M., Lapan, P. M., Wright, C. V. E. & Hogan, B. L. M. *Development* **115**, 639–647 (1992).

- Kao, K. R. & Elinson, R. P. *Dev. Biol.* **127**, 64–77 (1988).
- Cho, K. W. Y., Blumberg, B., Steinbeisser, H. & De Robertis, E. M. *Cell* **67**, 1111–1120 (1991).
- Sato, S. M. & Sargent, T. D. *Development* **112**, 747–753 (1991).
- Siegfried, E., Wilder, E. L. & Perrimon, N. *Nature* **367**, 76–79 (1994).
- Noordermeer, J., Klingensmith, J., Perrimon, N. & Nusse, R. *Nature* **367**, 80–82 (1994).
- McCrear, P. D., Brieher, W. M. & Gumbiner, B. M. *J. cell Biol.* **123**, 477–484 (1993).
- Heasman, J. *et al. Cell* **79**, 791–803 (1994).
- Dale, L., Matthews, G. & Colman, A. *EMBO J.* **12**, 4471–4480 (1993).
- Whitman, M. & Melton, D. A. *Nature* **357**, 252–254 (1992).
- Campbell, G., Weaver, T. & Tomlinson, A. *Cell* **74**, 1113–1123 (1993).
- Suzuki, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 10255–10259 (1994).
- Graff, J. M., Thies, R. S., Song, J. J., Celeste, A. J. & Melton, D. A. *Cell* **79**, 169–179 (1994).
- Bohn, H. *J. Embryol. exp. Morph.* **28**, 185–208 (1992).
- Struhl, G. & Basler, K. *Cell* **72**, 527–540 (1993).
- Diaz-Benjumea, F. J. & Cohen, S. M. *Development* **120**, 1661–1670 (1994).
- Harwood, A. J., Plyte, S. E., Woodgett, J., Strutt, H. & Kay, R. R. *Cell* **80**, 139–148 (1995).
- Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057–7070 (1984).
- Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488–492 (1985).
- Nieuwkoop, P. D. & Faber, J. *Normal table of Xenopus laevis* (Daudin) (North-Holland, Amsterdam, 1967).
- Nishimatsu, S., Suzuki, A., Shoda, A., Murakami, K. & Ueno, N. *Biochem. biophys. Res. Commun.* **186**, 1487–1495 (1992).
- Harland, R. M. *Meth. Cell Biol.* **36**, 685–695 (1991).
- Turner, D. L. & Weintraub, H. *Genes Dev.* **8**, 1434–1447 (1994).

ACKNOWLEDGEMENTS. We thank M. Taira, M. Rebagliati, P. Schwartzberg and L. Godley for suggestions and comments on the manuscript, M. R. for egg cDNA, D. Turner for pCS2 + n β -gal plasmid and N. Ueno for the XBMP-4 construct.

Explaining the spectrum of Sagittarius A* with a model of an accreting black hole

Ramesh Narayan, Insu Yi & Rohan Mahadevan

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

THE radio source at the centre of our Galaxy¹, Sagittarius A* (Sgr A*), seems to be a low-luminosity version of active galactic nuclei—a massive black hole that is accreting gas from the surrounding region^{1,2}. This idea is supported by observations of the gas and stars within 1 pc of Sgr A*, which appear to move under the influence of a large central mass^{1,3,4}. A recent determination of the upper limit^{5,6} to the hard X-ray emission from the Galactic Centre has posed a problem for this picture, however, as the mass accretion rate implied by applying a standard accretion model to the X-ray data is far below that estimated from the observations of gas flows. Here we present a new model of accretion onto Sgr A*, in which most of the energy released is carried along with the gas and lost into the black hole of mass $\sim 7 \times 10^5$ solar masses, rather than appearing as radiation. The model fits the observed spectrum of Sgr A* from radio to hard X-ray wavelengths, and reconciles the low observed luminosity with a high mass-accretion rate.

Figure 1 shows the spectral data on Sgr A*. The bulk of the emission occurs in the submillimetre band at frequency $\nu \approx 10^{12}$ Hz and in the near-infrared at $\nu \approx 10^{14}$ – $10^{14.5}$ Hz, but between these two bands there is a pronounced dip in the luminosity in the far-infrared band. Sgr A* also has highly variable emission in 1–30 keV X-rays. There is no direct information at optical and ultraviolet wavelengths because of absorption by the Galaxy¹; however, the maximum luminosity in these bands is limited^{1,7} to $< 3 \times 10^{39}$ erg s⁻¹. Further, because the Galactic Centre region is very crowded, the two near-infrared measurements should perhaps be considered as upper bounds.

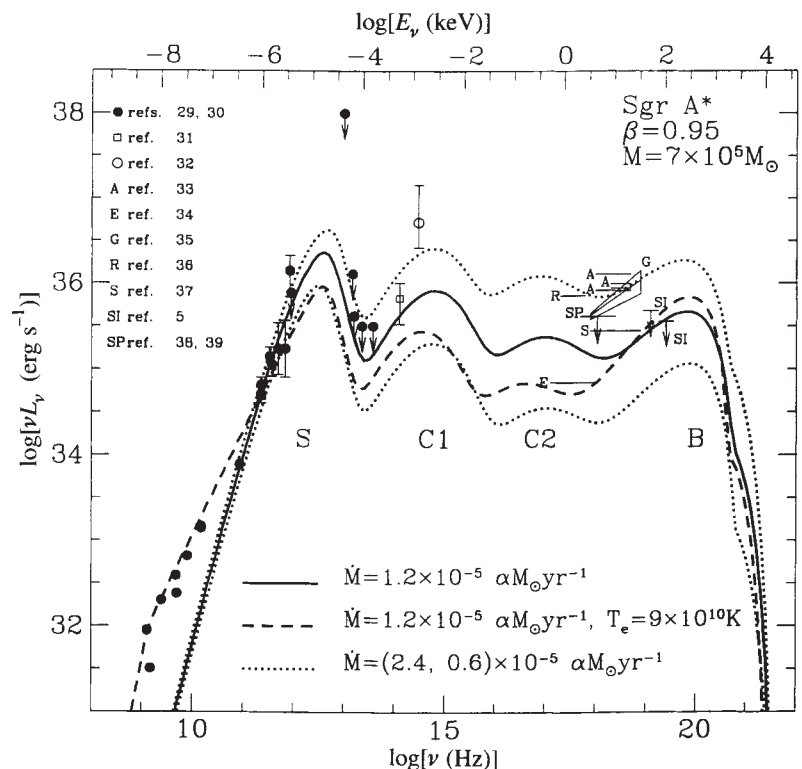
If Sgr A* is to be understood as an accretion-powered system, it is necessary to model the viscous flow of the accreting gas as well as the mechanisms whereby the dissipated energy is converted into the radiation we observe. The usual approach is to use the thin accretion disk model⁸ which assumes that cooling is efficient and that the energy released through viscosity is radiated immediately. But this model does not give a good fit to the observed spectrum of Sgr A*, and it is necessary to hypothesize a second component in the source with completely different radiation properties if one wishes to explain the observations⁹.

Some authors have expanded the thin-disk model to include the effects of radial pressure gradients and radial energy transport. This extended theory is referred to as the slim-disk model¹⁰ and one of its features is that it allows energy advection, whereby the gas transports a fraction f of the viscously dissipated energy as stored entropy and radiates only a fraction $1-f$; the fraction f is solved for self-consistently as a function of radius. The slim-disk equations have been used successfully in the study of disk instabilities^{10,11} and in modelling boundary layers of accretion disks around white dwarfs and pre-main-sequence stars^{12,13}.

The present work grew out of investigations of advection-dominated flows^{14–18}, where energy advection is not just a perturbation, or included merely for self-consistency, but where advection actually dominates the physics of the flow. In terms of the parameter f , these flows have a local radiative efficiency $1-f \ll 1$. Optically thin advection-dominated flows can occur^{16,18} around black holes for $\dot{M} < (10^{-4}–10^{-3})(M/10^6 M_\odot) M_\odot \text{ yr}^{-1}$, where M is the mass of the black hole, and $\dot{M}$ is the mass accretion rate.

The basic framework of our advection-dominated models is described in ref. 16. The accreting gas is considered to be a two-temperature plasma^{18,19} with the ions essentially at virial temperature at all radii and with the electrons significantly cooler than the ions in the inner regions. The two-temperature feature, plus the proper inclusion of advection and angular momentum transport in the accretion dynamics, distinguish our model from previous work. In view of the evidence for near-keplerian motions close to the Galactic Centre¹ we feel that it is important to consider rotating flows rather than the radial flows treated in some earlier work²⁰.

FIG. 1 The spectrum of Sgr A*. The horizontal axis gives the log of the frequency ν of the radiation along the bottom and the log of the photon energy E_ν along the top. The vertical axis gives νL_ν , where L_ν is the luminosity of the source between frequencies ν and $\nu + d\nu$. The symbols and horizontal lines represent various measurements of the spectrum of Sgr A*, taken from the references indicated. The measured fluxes were converted to luminosities assuming a distance of 8.5 kpc to the Galactic Centre. Arrows represent upper limits, and the sloping 'box' gives the range of fluxes and spectral slopes according to ref. 35. The thick solid curve shows the spectrum predicted by our best model, which has a black-hole mass of $(7 \times 10^5) M_\odot$ and a mass accretion rate of $\dot{M} = (1.2 \times 10^{-5}) \alpha M_\odot \text{ yr}^{-1}$. The four peaks, S, C1, C2 and B, are explained in the text. The upper and lower dotted curves refer to two models with $\dot{M} = (0.6, 2.4) \times 10^{-5} \alpha M_\odot \text{ yr}^{-1}$ respectively. The two models produce a factor of 30 variation in the X-ray flux. The dashed line is an isothermal model where we have arbitrarily assumed that the electrons are at a constant temperature of 9×10^9 K out to a radius of 3,000 black-hole radii.



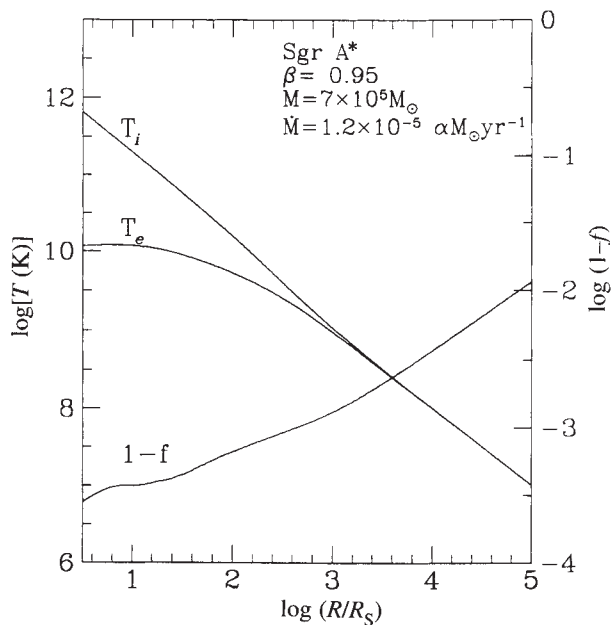


FIG. 2 Variation with radius R of the ion temperature, T_i , the electron temperature, T_e , and the radiative efficiency, $1-f$, in the model shown as a solid line in Fig. 1. (R_s is the black-hole radius.)

Our model has four parameters: (1) α , which describes the viscosity of the gas²¹; (2) β , which is defined such that $1-\beta$ is the fraction of the pressure contributed by magnetic fields; (3) M ; and (4) $\dot{M}$. We fix $\beta=0.95$ in all our models, corresponding to a magnetic pressure equal to one-tenth of equipartition. As the predicted spectrum depends only on the ratio $\dot{M}/\alpha$ rather than on $\dot{M}$ or α individually, we thus have only two adjustable parameters, M and $\dot{M}/\alpha$; α probably lies in the range 0.01–1.

The solid line in Fig. 1 shows the spectrum corresponding to our optimal model, with $M=(7 \times 10^5) M_\odot$ and $\dot{M}/\alpha=(1.2 \times 10^{-5}) M_\odot \text{ yr}^{-1}=(8 \times 10^{20}) \text{ g s}^{-1}$. This model satisfies most of the flux measurements and upper limits available on Sgr A*.

Figure 2 shows the variations of the ion temperature T_i , the electron temperature T_e , and the radiative efficiency $1-f$, as functions of radius R in this model. At large radii, T_i and T_e are both very nearly equal to the local virial temperature T_{vir} . Energy transfer from ions to electrons via Coulomb scattering is very efficient and so the two species come into thermal equilibrium. But the cooling of the electrons is inefficient, and the flow is therefore advection-dominated. For $R < 100 R_s$, where $R_s = 2.1 \times 10^{11} \text{ cm}$ is the size of the black hole, the electron temperature saturates at $T_e \approx 10^{10} \text{ K}$ but the ion temperature continues to track T_{vir} . Although electron cooling is efficient here, the ion–electron coupling becomes weak, and once again the flow is advection-dominated. Thus, at all radii, the radiative efficiency $1-f \ll 1$.

Figure 1 shows that there are four peaks in the model spectrum. The peak labelled S is due to synchrotron emission by the thermal electrons in the magnetic field of the plasma, the peaks C1 and C2 are due to Compton scattering (single scattering and double scattering, respectively) of the synchrotron radiation by the hot electrons, and the peak B is due to bremsstrahlung emission augmented by triply Compton-scattered photons. The positions at which the four peaks appear and their relative heights depend on the parameters of the model.

The position of the synchrotron peak S depends primarily on M , whereas its height varies approximately as $\dot{M}/\alpha$. Thus, merely by fitting the submillimetre-wave data and requiring that the model satisfy the far-infrared limits at $\nu \approx 10^{13}$ – 10^{14} Hz , it is possible to estimate M and $\dot{M}$ to within a factor of ~ 2 . The sharp drop of the emission at the high- ν end of the S peak is a

characteristic signature of self-absorbed synchrotron emission from a thermal plasma²². It can also be explained with a single-energy population of electrons²³. However, models that invoke a power-law distribution of relativistic electrons will not be consistent with the dip in the spectrum.

Because the synchrotron emission is highly self-absorbed, the emission at each ν in the S peak originates essentially at a single radius. The radiation at 86 GHz, for instance, comes from a radius of $1.2 \times 10^{13} \text{ cm}$ in our model (see also ref. 24). This is consistent with recent measurements^{25,26} which give an upper limit of $1.6 \times 10^{13} \text{ cm}$ for the size of the source.

At radio frequencies, $\nu < 10^{11} \text{ Hz}$, our model spectrum lies below the observations. The emission at these frequencies is from material at several $100 R_s$, where our model gives relatively low electron temperatures. However, it is possible to envisage non-local transport processes from the interior, such as streaming of suprathermal electrons, which may be able to enhance the effective temperature out here and thereby increase the emission. As an example, the dashed curve in Fig. 1 shows a model where we have arbitrarily assumed a constant temperature of $T_e = 9 \times 10^9 \text{ K}$ out to $R = 3,000 R_s$. The non-local energy transport could well be irregular and may lead to variability²⁷ in the source below 10^{11} Hz .

The frequency shift between the S and C1 peaks, and that between C1 and C2, depend on the mean energy boost $\langle A \rangle$ of a photon per Compton scattering. This depends on T_e . In the self-consistent models shown here, the maximum electron temperature is $\sim (1.1\text{--}1.3) \times 10^{10} \text{ K}$ (Fig. 2), which gives $\langle A \rangle \approx 60\text{--}90$. Note that the electron scattering optical depth is very small, $\tau_{\text{es}} < 10^{-2}$. Therefore, most of the synchrotron photons escape and only a small fraction ($\sim \tau_{\text{es}}$) is Compton-scattered; the fraction that is scattered twice is $\sim \tau_{\text{es}}^2$. The heights of the C1 and C2 peaks relative to S are therefore proportional to $\langle A \rangle \tau_{\text{es}}$ and $\langle A \rangle^2 \tau_{\text{es}}^2$, respectively. As the flux in the S peak and the value of τ_{es} are both $\propto \dot{M}$, the flux in the C1 peak varies as $\sim \dot{M}^2$ and that in C2 as $\sim \dot{M}^3$.

As the data in Fig. 1 show, the X-ray flux of Sgr A* at energies $\sim 1\text{--}30 \text{ keV}$ is variable by almost a factor of 20. The large variability is quite naturally explained in our model as due to modest changes in $\dot{M}$ as a function of time. The dotted curves in Fig. 1 show two models with $\dot{M}/\alpha$ greater than, and smaller than, that of the optimal model by a factor of 2. Even with such a small variation in $\dot{M}$, the two models more than cover the range of fluxes observed in X-rays. For the same change in $\dot{M}$, the flux at $10^{11}\text{--}10^{12} \text{ Hz}$ hardly varies at all, while the submillimetre and near-infrared signals in the S and C1 peaks show lower-amplitude variations than in X-rays. Because the S, C1 and C2 peaks are closely coupled in the model, fluctuations in their heights should show a progression in amplitude and should be correlated in time, a testable prediction.

The bremsstrahlung peak B occurs at the thermal frequency, $\nu = kT_e/h$, so its position is determined primarily by T_e . The data in Fig. 1 suggest that this peak may be too close to the hard X-ray limits. With a higher value of T_e , as might arise from a more efficient electron–ion coupling²⁸, the peak can be moved to a higher ν without significantly affecting the other peaks in the spectrum. Observations of the γ -rays emitted from Sgr A* at energies $\sim 1 \text{ MeV}$ will be able to constrain the properties of the bremsstrahlung emission, and thereby help determine T_e .

We should emphasize that there is very little fine-tuning in our model. Merely by fitting the S peak we are able to estimate M and $\dot{M}/\alpha$, and with these parameters we obtain a more-or-less satisfactory fit to the rest of the spectrum. The dip between S and C1 is a robust property of thermal synchrotron emission²² which nicely explains a feature in the data that is very difficult to fit with other models. Furthermore, all the emission comes from the same hot electrons located within $R < (10\text{--}100) R_s$, and the different peaks in the spectrum are just signatures of different radiation mechanisms. Considering how simple the model is in

its basic structure, the level of agreement with the data is quite good.

In our model of Sgr A*, more than 99.9% of the energy is advected with the gas and is carried into the black hole through the event horizon, while less than 0.1% is radiated (see $1-f$ in Fig. 2). Any model of Sgr A* that fails to include advection and energy loss through a horizon will necessarily have to invoke a very small $\dot{M}$ ($\sim 10^{17} \text{ g s}^{-1}$), because otherwise the predicted luminosity ($\sim 0.1 \dot{M} c^2$) will greatly exceed that observed^{5,7}. Such a low $\dot{M}$ is in serious conflict with an independent estimate of $\dot{M}$ ($\sim 10^{22} \text{ g s}^{-1}$) derived²⁰ from observations of gas flows near the Galactic Centre. Even our advection-dominated model, with $\alpha = 1$ and $\dot{M} = 8 \times 10^{20} \text{ g s}^{-1}$, falls short of this estimate by a factor of ~ 13 , but may be considered marginally consistent given the uncertainties. $\square$

Received 4 November 1994; accepted 2 March 1995.

- Genzel, R. & Townes, C. H. A. *Rev. Astr. Astrophys.* **25**, 377–423 (1987).
- Lynden-Bell, D. & Rees, M. *Mon. Not. R. astr. Soc.* **152**, 461–475 (1971).
- Lacy, J. H., Achtermann, J. M. & Serabyn, E. *Astrophys. J.* **380**, L71–L74 (1991).
- McGinn, J. T., Sellgren, K., Becklin, E. E. & Hall, D. N. B. *Astrophys. J.* **338**, 824–840 (1989).
- Goldwurm, A. et al. *Nature* **371**, 589–591 (1994).
- Grindlay, J. E. *Nature* **371**, 561–562 (1994).
- Falcke, H., Biermann, P. L., Duschl, W. J. & Mezger, P. G. *Astr. Astrophys.* **270**, 102–106 (1993).
- Frank, J., King, A. & Raine, D. *Accretion Power in Astrophysics* (Cambridge Univ. Press, 1992).
- Falcke, H., Mannheim, K. & Biermann, P. L. *Astr. Astrophys.* **278**, L1–L4 (1993).

- Abramowicz, M., Czerny, B., Lasota, J. P. & Szuszkiewicz, E. *Astrophys. J.* **332**, 646–658 (1988).
- Honma, F., Matsumoto, R. & Kato, S. *Publs astr. Soc. Japan* **43**, 147–168 (1991).
- Narayan, R. & Popham, R. *Nature* **362**, 820–822 (1993).
- Popham, R., Narayan, R., Hartmann, L. & Kenyon, S. *Astrophys. J.* **415**, L127–L130 (1993).
- Narayan, R. & Yi, I. *Astrophys. J.* **428**, L13–L16 (1994).
- Narayan, R. & Yi, I. *Astrophys. J.* **444**, 231–243 (1995).
- Narayan, R. & Yi, I. *Astrophys. J.* (in the press).
- Abramowicz, M., Chen, X., Kato, S., Lasota, J. P. & Regev, O. *Astrophys. J.* **438**, L37–L40 (1995).
- Rees, M. J., Begelman, M. C., Blandford, R. D. & Phinney, E. S. *Nature* **295**, 17–21 (1982).
- Shapiro, S. L., Lightman, A. P. & Eardley, D. M. *Astrophys. J.* **204**, 187–199 (1976).
- Melia, F. *Astrophys. J.* **387**, L25–L28 (1992).
- Shakura, N. I. & Sunyaev, R. A. *Astr. Astrophys.* **24**, 337–355 (1973).
- Melia, F. *Astrophys. J.* **426**, 577–585 (1994).
- Duschl, W. J. & Lesch, H. *Astr. Astrophys.* **286**, 431–436 (1994).
- Melia, F., Jokipii, J. R. & Narayanan, A. *Astrophys. J.* **395**, L87–L90 (1992).
- Rogers, A. E. E. et al. *Astrophys. J.* **434**, L59–L62 (1994).
- Backer, D. C. et al. *Science* **262**, 1414–1416 (1993).
- Zhao, J.-H., Ekers, R. D., Goss, W. M., Lo, K. Y. & Narayan, R. in *The Center of the Galaxy* (ed. Morris, M.) 535–538 (IAU Symp. No. 136, Kluwer, Dordrecht, 1989).
- Begelman, M. C. & Chiueh, T. *Astrophys. J.* **332**, 872–890 (1988).
- Zylka, R., Mezger, P. G. & Lesch, H. *Astr. Astrophys.* **261**, 119–129 (1992).
- Zylka, R., Mezger, P. G., Ward-Thompson, D., Duschl, W. J. & Lesch, H. *Astr. Astrophys.* (in the press).
- Eckart, A. et al. *Nature* **355**, 526–529 (1992).
- Rees, M. J., Zinnecker, H., Moneti, A. & Melnick, J. *Astr. Astrophys.* **257**, 515–522 (1991).
- Pavlinkii, M. N., Grebenev, S. A. & Sunyaev, R. A. *Soviet Astr. Lett.* **18**, 116–120 (1992).
- Watson, M. G., Willingale, R., Grindlay, J. E. & Hertz, P. *Astrophys. J.* **250**, 142–154 (1981).
- Pavlinkii, M. N., Grebenev, S. A. & Sunyaev, R. A. *Astrophys. J.* **425**, 110–121 (1994).
- Predehl, P. & Trümper, J. *Astr. Astrophys.* **290**, L29–L32 (1994).
- Skinner, G. K. et al. *Nature* **330**, 544–547 (1987).
- Kawai, N. et al. *Astrophys. J.* **330**, 130–141 (1988).
- Skinner, G. K. in *The Center of the Galaxy* (ed. Morris, M.) 567–580 (IAU Symp. No. 136, Kluwer, Dordrecht, 1989).

ACKNOWLEDGEMENTS. We thank D. Barret, W. Duschl, H. Falcke, J. Grindlay, and the referees, F. Melia and G. Skinner, for useful comments.

Large third-order optical nonlinearities in transition-metal oxides

Masanori Ando^{*†}, Kohei Kadono^{*},
Masatake Haruta^{*}, Toru Sakaguchi^{*}
& Masaru Miya^{*}

^{*} Osaka National Research Institute, AIST, Midorigaoka 1, Ikeda, Osaka 563, Japan

[†] Light and Material Group, PRESTO, Research Development Corporation of Japan (JRDC), Midorigaoka 1, Ikeda, Osaka 563, Japan

ADVANCES in the field of optical computing^{1–3} will require the development of materials that combine a large nonlinear optical response with a fast response time. For many applications, this translates into a third-order nonlinear optical susceptibility, $\chi^{(3)}$, in excess of 10^{-8} e.s.u., and a response time faster than 10 ps (ref. 4). Although a wide range of inorganic^{5–18} and organic^{19–21} materials have been found to exhibit a large $\chi^{(3)}$, either the response times tend to be far too slow or the materials are not sufficiently stable for device applications. Recently, the transition-metal oxide Fe_2O_3 was found to have a large $\chi^{(3)}$ (ref. 22). Here we show that oxides of several other 3d transition metals show a similarly large nonlinear optical response; moreover, we find that a significant contribution to the overall $\chi^{(3)}$ ($\sim 10^{-8}$ e.s.u. in the case of V_2O_5) has a response time of the order of 35 ps.

Thin film samples of V_2O_5 , Cr_2O_3 , Mn_3O_4 , Fe_2O_3 , Co_3O_4 , CuO and NiO , all with centrosymmetric crystal structures, were prepared by thermal decomposition at 380 °C in air of metal alkylcarboxylates spin-coated on a glass-plate substrate ($18 \times 18 \times 0.1 \text{ mm}$)²³. The thickness (d) and refractive index (n) of the oxide films were measured by using a Mizojiri DHA-XA2 ellipsometer at a wavelength of 633 nm (see Table 1). Optical absorption in the visible/near-infrared region was measured at room temperature with a Shimadzu UV-3100PC double-beam spectrometer.

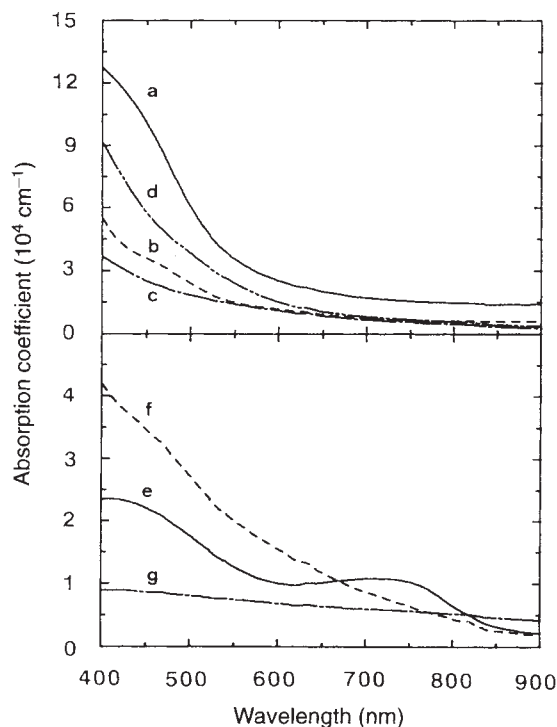


FIG. 1 Electronic absorption spectra of 3d transition-metal oxide films. a, V_2O_5 ; b, Cr_2O_3 ; c, Mn_3O_4 ; d, Fe_2O_3 ; e, Co_3O_4 ; f, CuO and g, NiO .

Figure 1 shows the optical absorption, in units of absorption coefficient (α), for each oxide film. Because the photon energy of ultraviolet light corresponds to the energy gap of semiconducting oxides, the absorption of these oxide films increases as the wavelength approaches the ultraviolet region.

The value of $\chi^{(3)}$ was evaluated by the phase-conjugation-type degenerate four-wave mixing method²⁴, in which one of the pump beams and the probe beam form a phase or intensity

TABLE 1 Optical properties of some third-order nonlinear optical materials

Material	<i>d</i> (nm)	<i>n</i>	α (10^4 cm^{-1})	$\chi^{(3)*}$ (10^{-8} e.s.u.)	$\chi^{(3)*}/\alpha$ ($10^{-12} \text{ e.s.u. cm}$)	$\chi^{(3)\dagger}$ (10^{-9} e.s.u.)	$\chi^{(3)\dagger}/\alpha$ ($10^{-13} \text{ e.s.u. cm}$)	$\chi^{(3)*}/\chi^{(3)\dagger}$
V ₂ O ₅	33.5	2.37	4.22	31	7.3	12	2.7	27
Cr ₂ O ₃	44.0	1.23	1.72	1.9	1.1	0.75	0.43	25
Mn ₃ O ₄	32.2	1.78	1.55	2.8	1.8	2.0	1.3	14
Fe ₂ O ₃	24.7	2.02	2.85	5.6	2.0	4.2	1.5	13
Co ₃ O ₄	60.4	1.71	1.42	4.6	3.2	4.5	3.2	10
CuO	33.3	1.87	2.22	5.7	2.6	4.7	2.1	12
NiO	85.0	1.72	0.765	(<0.01) [‡]				
Li-NiO§	79.8	1.74	0.652	0.14	0.22			
Au-glass	190	1.5	3.74	12	3.2			
PBTABQ¶		1.762	8.7			45	5.2	

Optical properties of 3d transition-metal oxide films and other representative third-order nonlinear optical materials. Symbols used; *d*, film thickness; *n*, refractive index; α , absorption coefficient.

* Determined using 7-ns-pulse laser.

† Determined using 35-ps-pulse laser.

§ NiO film doped with Li (atomic ratio Li/Ni = 5/100).

‡ Lower than the detection limit.

|| Glass doped with ultra-fine Au particles (Au content, 6.3 at.% (40 wt%)). The value of $\chi^{(3)}$ was estimated at a wavelength of 532 nm using a 7-ns-pulse laser for excitation (refs 9–11).

¶ Poly (α -[5,5'-bithiophenediyl]p-acetoxybenzylidene-block- α -[5,5'-bithiophenequinodimethanediyl]). The $\chi^{(3)}$ value for the solid phase shown here was deduced from the value for the sulphuric-acid solution which was estimated at a wavelength of 532 nm using a 25-ps-pulse laser for excitation²¹. The values of α and $\chi^{(3)}/\alpha$ were calculated from the data described in ref. 21.

grating that modulates the refractive index of a material via the third-order susceptibility. The other pump beam is diffracted by the grating to produce the phase-conjugate beam which is detected using a streak camera^{9–11}. The phase-conjugate reflectivity, defined as the intensity ratio of the phase-conjugate beam to the probe beam, is related to the value of $\chi^{(3)}$ by a known formula^{9,10}. For these studies, the light source was a frequency-doubled Q-switched Nd:YAG laser with a wavelength of 532 nm and pulse duration of 7 ns or 35 ps. The beam diameter was 0.3 cm at the sample position. The applied peak power density was varied from 0.1 to 1.3 MW cm⁻² for the 7-ns-pulse laser and was held constant at 55 MW cm⁻² for the 35-ps-pulse laser. As a calibration standard, CS₂ was used^{9–11}. The contribution of the glass substrate when evaluating $\chi^{(3)}$ of the transition-metal oxide films was negligible.

The typical dependence of the phase-conjugate reflectivity on the pumping intensity measured for a V₂O₅ film using the 7-ns-pulse laser is shown in Fig. 2. Linear correlation was found in the log-log plot with a slope of ~ 2 , which is evidence for the third-order nonlinear optical effect⁹. The third-order nonlinearity does not saturate up to a pumping intensity of at least 1.3 MW cm⁻². The absence of saturation is advantageous, and

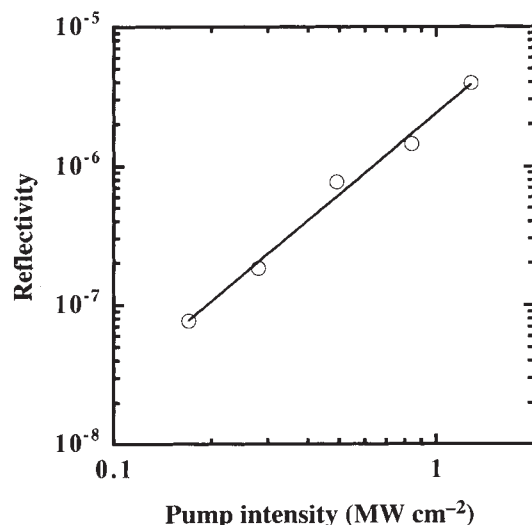


FIG. 2 Intensity of degenerate four-wave mixing signal from a V₂O₅ film in units of reflectivity as a function of mean pump intensity.

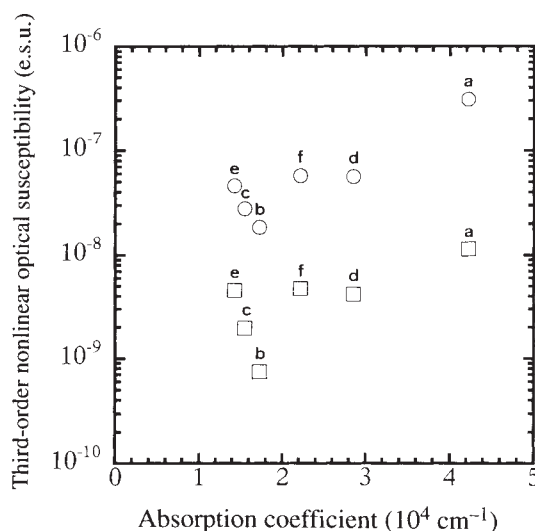


FIG. 3 Relationship between the third-order nonlinear optical susceptibility and the absorption coefficient of oxide films. $\circ$, using 7-ns-pulse laser; $\square$, using 35-ps-pulse laser. a, V₂O₅; b, Cr₂O₃; c, Mn₃O₄; d, Fe₂O₃; e, Co₃O₄ and f, CuO.

distinguishes this compound from most other nonlinear semiconductors which saturate at lower pump intensities²⁵.

Table 1 summarizes the linear and nonlinear optical properties of the oxide-film samples, and compares these nonlinearities with those of other representative third-order nonlinear optical materials. All oxides (except for NiO) were found to have large values of $\chi^{(3)}$, within the range 10^{-8} – 10^{-7} e.s.u. In particular, V₂O₅ gave the highest value of $\chi^{(3)}$, namely 3.1×10^{-7} e.s.u. It should be noted that the $\chi^{(3)}$ of V₂O₅ compares favourably with those of other representative third-order nonlinear materials, such as high density Au-doped glass^{9–11} and organic polymer²¹ measured under similar conditions.

The resulting $\chi^{(3)}$ ranking of oxides is shown in Table 1 and Fig. 3; the absorption spectra (Fig. 1) of oxides indicate that $\chi^{(3)}$ generally increases with increasing α . Therefore, we speculate that the strong nonlinearity of the oxides is due to the following absorptive nonlinear effects. First, the band-filling effect^{26,27} which originates from the change of refractive index brought about by an increase in carrier density in the excited state. This is caused by absorbing intense light with photon energies near

the energy gap of semiconductors. Second, the thermal effect⁸ which results from the change of refractive index caused by an increase in temperature on interaction with the laser beam. As the thermal nonlinear process takes place in a period longer than several tens of picoseconds, the band-filling effect and the thermal effect may be responsible for the fast $\chi^{(3)}$ component observed using the 35-ps-pulse laser and the slower $\chi^{(3)}$ component observed using the 7-ns-pulse laser, respectively. The nonlinear optical properties of a material may be characterized by a figure of merit, $\chi^{(3)}/\alpha$ which is related to the inherent strength of the nonlinear process. Vanadium oxide has advantages over other 3d transition-metal oxides, and also compares favourably with other representative third-order nonlinear materials, when considered from the viewpoint of $\chi^{(3)}/\alpha$ (Table 1).

The $\chi^{(3)}$ obtained using the 35-ps-pulse laser is 10–30 times smaller than that obtained using the laser with 7-ns pulses. This factor (10–30) is a measure of the relative contributions of the slower and faster $\chi^{(3)}$ components, and is found to be smallest in Co_3O_4 . The large contribution from the faster $\chi^{(3)}$ component for Co_3O_4 might be related to the relatively high electrical conductivity of Co_3O_4 (compared to the other oxides investigated here), because high density and mobility of the charge carrier in semiconductors has been reported to enhance the nonlinear optical effect^{28,29}. The considerable contribution of the fast $\chi^{(3)}$ component (with picosecond-timescale response) in the 3d transition-metal oxides investigated here give these compounds significant advantages over other well known third-order nonlinear oxides such as LiNbO_3 (refs 16, 17) and BaTiO_3 (refs 17, 18), the response times of which are in the range 10^{-3} –1 s.

We are working on 3d transition-metal oxides, doped with electron donors or with electron acceptors, where the value of $\chi^{(3)}$ is expected to be enhanced by several orders of magnitude as compared with undoped oxides. For instance, we find that

even NiO , with $\chi^{(3)} < 10^{-10}$ e.s.u. in the undoped state, shows an increase in $\chi^{(3)}$ of at least 1–2 orders of magnitude when doped with a small amount of Li; an (unwanted) increase in α is not observed (Table 1). We assume that the third-order optical nonlinearities of other oxides would be also enhanced by doping, possibly leading to nonlinearity sufficient for practical optoelectronic device applications. $\square$

Received 17 October 1994; accepted 7 March 1995.

1. Eaton, D. F. *Science* **253**, 281–287 (1991).
2. Kwong, S., Rakuljic, G. A. & Yariv, A. *Appl. Phys. Lett.* **48**, 201–203 (1986).
3. Soffer, B. H., Dunning, G. J., Owechko, Y. & Marom, E. *Opt. Lett.* **11**, 118–120 (1986).
4. Nakanishi, H., Matsuda, H. & Okada, S. in *Optical Functional Materials* 41–105 (Soc. of Polymer Science, Japan, Kyoritsu, Tokyo, 1991).
5. Nasu, H., Ibara, Y. & Kubodera, K. *J. non-cryst. Solids* **110**, 229–234 (1989).
6. Jain, R. K. & Lind, R. C. *J. opt. Soc. Am.* **73**, 647–653 (1983).
7. Justus, B. L., Seaver, M. E., Ruller, J. A. & Campillo, A. *J. Appl. Phys. Lett.* **57**, 1381–1383 (1990).
8. Hache, F., Ricard, D., Flytzanis, C. & Kreibig, U. *Appl. Phys. A47*, 347–357 (1988).
9. Fukumi, K. *et al. J. appl. Phys.* **30**, L742–L744 (1991).
10. Fukumi, K. *et al. J. appl. Phys.* **75**, 3075–3080 (1994).
11. Kadono, K. *et al. Mater. Res. Soc. Symp. Proc.* **283**, 903–908 (1993).
12. Magruder, R. H. *et al. Appl. Phys. Lett.* **62**, 1730–1732 (1993).
13. Akai, T. *et al. J. Ceram. Soc. Jap.* **101**, 105–107 (1993).
14. Magruder, R. H., Haglund, R. F. Jr, Yang, L., Wittig, J. E. & Zuhur, R. A. *J. appl. Phys.* **76**, 708–715 (1994).
15. Gibbs, H. M. *et al. Appl. Phys. Lett.* **41**, 221–222 (1982).
16. Glass, A. M. *Opt. Engng* **17**, 470–479 (1978).
17. Günter, P. *Phys. Rep.* **93**, 199–299 (1982).
18. Valley, G. C. & Klein, M. B. *Opt. Engng* **22**, 704–711 (1983).
19. Sauteret, C. *et al. Phys. Rev. Lett.* **56**, 956–959 (1976).
20. Matsumoto, S., Kubodera, K., Kurihara, T. & Kaino, T. *Appl. Phys. Lett.* **51**, 1–2 (1987).
21. Jenekhe, S. A., Chen, W., Lo, S. & Flom, S. R. *Appl. Phys. Lett.* **57**, 126–128 (1990).
22. Hashimoto, T., Yoko, T. & Sakka, S. *J. Ceram. Soc. Jap.* **101**, 64–68 (1993).
23. Ando, M., Kobayashi, T. & Haruta, M. *J. chem. Soc., Faraday Trans.* **90**, 1011–1013 (1994).
24. Kobayashi, T., Terasake, A., Hattori, T. & Kurokawa, K. *Appl. Phys. B47*, 107–125 (1988).
25. Yuen, S. Y. & Becla, P. *Opt. Lett.* **8**, 356–358 (1983).
26. Wherrett, B. S. & Higgins, N. A. *Proc. R. Soc. A379*, 67–90 (1982).
27. Miller, D. A. B., Seaton, C. T., Prise, M. E. & Smith, S. D. *Phys. Rev. Lett.* **47**, 197–200 (1981).
28. Strait, J. & Glass, A. M. *J. opt. Soc. Am B3*, 342–344 (1986).
29. Bylsma, R. B., Bridenbaugh, P. M., Olson, D. H. & Glass, A. M. *Appl. Phys. Lett.* **51**, 889–891 (1987).

Fatigue-free ferroelectric capacitors with platinum electrodes

C. A-Paz de Araujo*[†], J. D. Cuchiaro*,
L. D. McMillan*, M. C. Scott* & J. F. Scott*[‡]

* Symetrix Corporation, 5055 Mark Dabbling Boulevard, Colorado Springs, Colorado 80918, USA

[†] Department of Electrical and Computer Engineering, University of Colorado, Colorado Springs, Colorado 80933-7150, USA

A SIGNIFICANT fraction of the computer memory industry is at present involved in the manufacture of non-volatile memory devices¹—that is, devices which retain information when power is interrupted. For such applications (and also for volatile memories), the use of capacitors constructed from ferroelectric thin films has stimulated much interest¹. In such structures, information is stored in the polarization state of the ferroelectric material itself, which should in principle lead to lower power requirements, faster access time and potentially lower cost¹. But the use of ferroelectrics is not without problems; the memories constructed to date have generally suffered from poor retention of stored information and degradation of performance ('fatigue') with use^{1–3}. Here we describe the preparation and characterization of thin-film capacitors using ferroelectric materials from a large family of layered perovskite oxides, exemplified by $\text{SrBi}_2\text{Ta}_2\text{O}_9$, $\text{SrBi}_2\text{NbTaO}_9$ and $\text{SrBi}_4\text{Ta}_4\text{O}_{15}$. The

structural flexibility of these materials allows their properties to be tailored so that many of the problems associated with previous ferroelectric memories are avoided. In particular, our capacitors do not show significant fatigue after 10^{12} switching cycles, and they exhibit good retention characteristics and low leakage currents even with films less than 100 nm thick.

Ferroelectrics subjected to repetitive polarization switching exhibit 'fatigue', a decrease in the amount of charge switched each time. For ferroelectric oxides, such as the commonly used perovskite titanates, configured with metal electrodes (usually platinum), fatigue is thought to arise from three different microscopic causes. (1) Stress relaxation^{2,3} of 90° domains in pseudocubic crystals (such as the barium or lead titanate-zirconate family); there the orthogonal domains slowly relax back to 180° configurations as mechanical stresses are released internally, and in doing so reorient and reduce the net polarization. The 'pinning' of 180° domains by stress or charged defects is a related fatigue mechanism. (2) 'Poling' of charged defect pairs^{3,6}, such as lead-vacancy, oxygen-vacancy neighbours; these dipoles become aligned with repetitive application of large electric fields and are very slow (taking seconds or longer) to reverse, thus reducing or cancelling some of the switched lattice polarization. (3) Space-charge accumulation at or near the electrode-ferroelectric interface^{7,10}; this compensates for the applied voltage and acts as a detrimental screening of external fields. Space-charge injection into the ferroelectric also leads to oxidation of the Pt electrode and valence conversion of metal ions such as Ti in the ferroelectric. (For a metal electrode with a 5-eV work function, an applied field of 10–100 MV m⁻¹ produces injected current¹¹ of several microamps per cm².) A related parameter is 'retention'—essentially the 'shelf-life' of stored information.

Two generic methods for minimizing this fatigue phenomenon have been considered. First, modification of the electrode. The

[‡] Permanent address: Faculty of Applied Science, Royal Melbourne Institute of Technology, Melbourne, Victoria 3001, Australia.

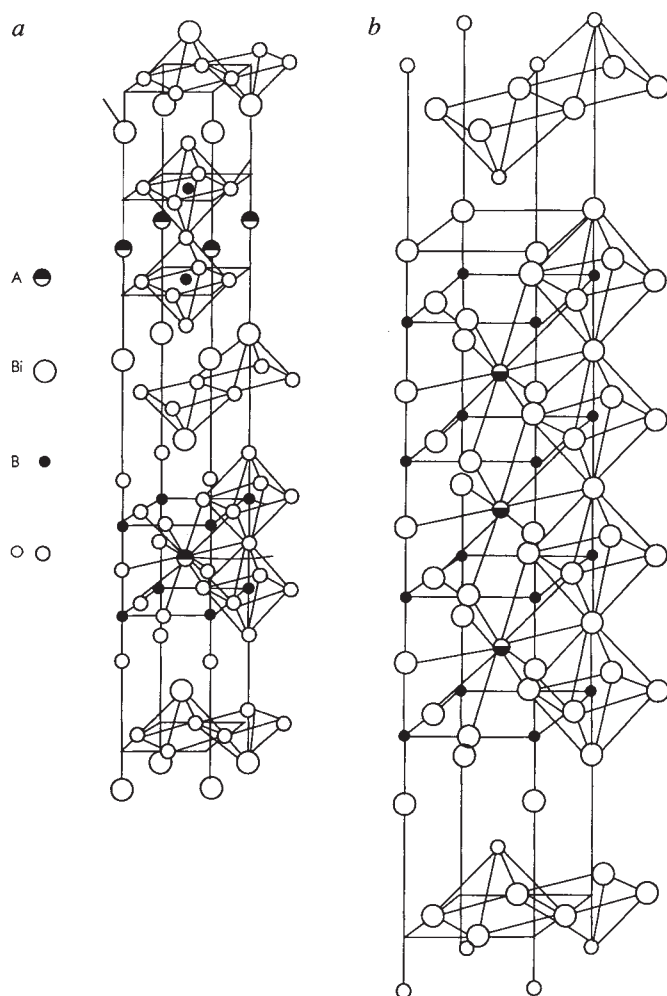


FIG. 1 *a*, Lattice structure of $ABi_2B_2O_9$ (where A is a divalent metal such as Sr, Ba, Pb; B is a metal of valence +5, usually Ta or Nb); *b*, lattice structure of $ABi_2B_4O_{15}$ (where A is a divalent metal such as Sr, Ba, Pb; B is a metal of valence +4, usually Ti).

use of iridium oxide is very well known in biomedical applications¹², where the interface between metal electrodes and tissue containing oxygen is a similar problem. Unlike platinum, iridium (or ruthenium) readily forms metastable oxides. These metastable oxides can reduce or re-oxidize reversibly and repeatedly without degradation, thus permitting the temporal dynamics of space-charge and oxygen-vacancy concentrations to be smoothed out or 'buffered' in a way that maintains a constant level of surface charging and electron injection from the electrode^{12,13}. The electrode minimizes large accumulations of free charge ('space charge') at the electrode-ferroelectric interface that might otherwise hinder retention or accelerate fatigue processes. This approach, solving the space-charge problem at the electrode interface by using an electrode material with a complex, dynamical electrochemical response, is used in lead zirconate-titanate (PZT) ferroelectric prototype memories with iridium/iridium oxide¹⁴ and with ruthenium oxide¹⁵⁻¹⁸. However, its disadvantage is that the resulting device is electrically 'leaky', with d.c. leakage currents that may be too high for some non-volatile memory devices and for dynamic random-access memories (DRAMs)^{19,20}.

The second scheme for minimizing fatigue is to use ordinary metal (for example, Pt) electrodes, but to modify the ferroelectric. Such modification can be done, for example, by donor doping or by employing a more complicated ferroelectric structure that will carry out the dynamic role described above of compensating in some way for space-charge and/or oxygen gradients near the electrode, thus reducing the possibility of excessive charge injection into the dielectric. Here we report the development of such ferroelectrics of the layered-structure perovskite oxide family²¹⁻²³ epitomized by $SrBi_2Ta_2O_9$, $SrBi_2NbTaO_9$, $SrBi_4Ti_4O_{15}$ and barium-based isomorphs. These materials consist of perovskite 'building blocks' of TaO_6 , NbO_6 or TiO_6 octahedra, separated at intervals by bismuth oxide planes (Fig. 1).

The preparation²⁴ of $SrBi_2Ta_2O_9$ or $SrBi_2NbTa_2O_9$ involves metal-organic decomposition of a solution of the precursors given below, spun on a metallized silicon substrate at 1,500 r.p.m. with a photo-resist spinner at 295 K, or alternatively, injected onto a rotating substrate at ambient temperatures in a liquid-phase chemical vapour deposition machine described elsewhere²⁵. The preferred process is to react tantalum butoxide and metallic strontium with 2-ethylhexanoic acid in xylene, followed by addition of bismuth 2-ethylhexanoate. The resulting

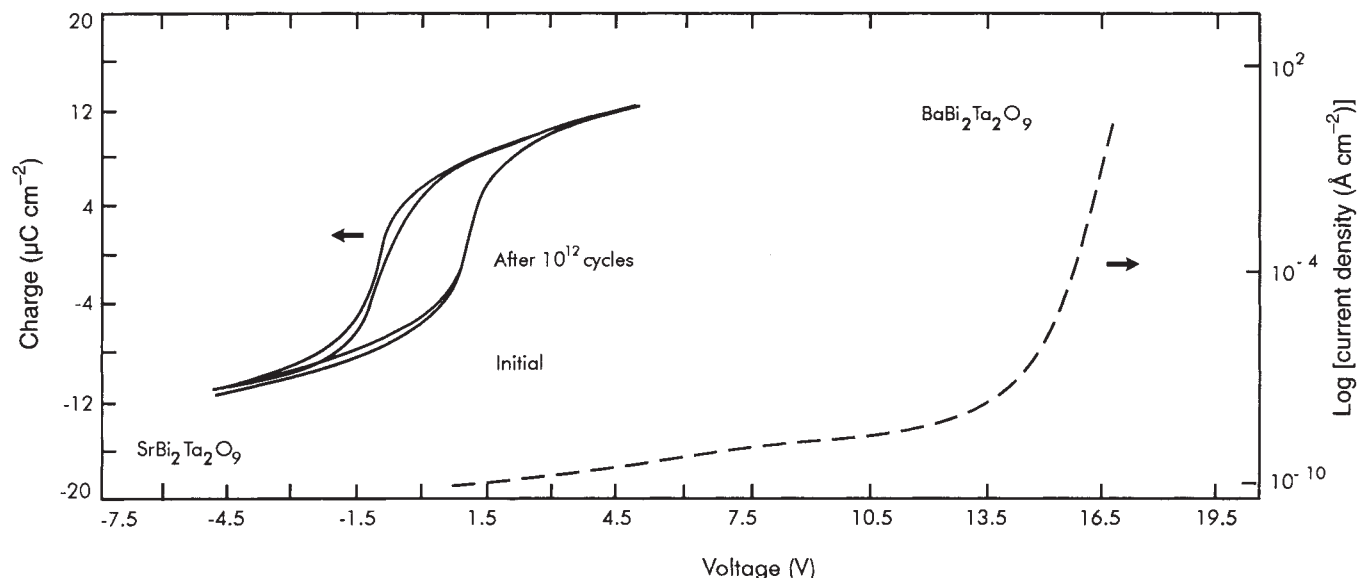


FIG. 2 Solid curves show hysteresis in $SrBi_2Ta_2O_9$. The two superimposed traces are the initial (virgin) trace and that after 10^{12} cycles at 280 kV cm^{-1} . Dashed curve, d.c. leakage current in a 240-nm-thick

film of $BaBi_2Ta_2O_9$ at 293 K. At the intended operating voltage of 5 V, the leakage is 1.0 nA cm^{-2} . Data were collected on a fast ramp of 10 kV cm^{-1} per s; slower ramp rates give even lower absolute values.

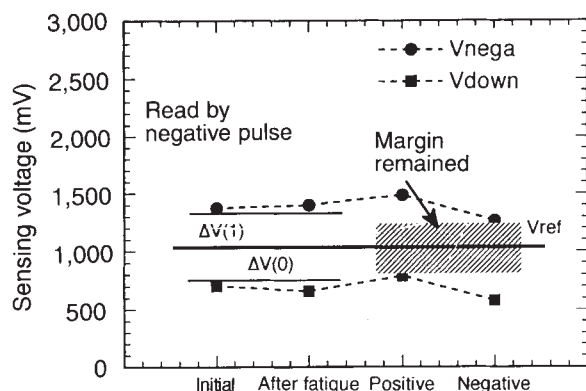


FIG. 3 Pulsed switching data on the specimen in Fig. 2a (10 nF capacitor) showing excellent retention characteristics. These measurements were made under the worst-case conditions of a thin film (175 nm), high temperature (85 °C), and after 10^9 polarization reversals. The notations refer to the voltage detected for polarization reversal (switching from 'up' to 'down'—a '1' in the boolean algebra of the memory) or no reversal (cell already 'down'—a '0' in the memory).

films are annealed by rapid thermal processing, with a 90-s anneal to a maximum temperature of 810 °C, or a furnace anneal at 700 °C for 60 min.

We chose these materials because the repeat distance of the bismuth oxide planes can be varied precisely by selection of the exact Bi/Ta ratio of precursors, the ambient pressure and deposition rate. Repetition distances of two, three, four or more perovskite ABO_3 'blocks' can be obtained between bismuth planes (where A is a divalent metal such as Sr; B is Nb, Ta or Ti); this can produce crystals of perfect translational symmetry and also incommensurate (modulated) structures. In general, as these bismuth oxide planes have a net electrical charge, their positioning in the lattice is self-regulated to compensate for space charge near the electrodes (free energy during the growth process is minimized by compensating for space charge). We can control the electrode interfacial space charge and hence the ferroelectric structure by choice of electrode work function, Bi/Ta precursor ratio, and processing parameters mentioned above (temperature, deposition rate and pressure). In this way some of the microscopic processes that would otherwise lead to fatigue are controlled within the ferroelectric itself, rather than in the electrode (as in the PZT/iridium oxide devices). The result is that ordinary metal (Pt) electrodes can be used. Fully functional 256-kb random-access memories (RAMs) have been manufactured with these materials²⁶.

Figure 2 illustrates the leakage currents and hysteresis curves that are obtained from materials of this family. Leakage currents as low as 1.0 nA cm^{-2} at 80-nm thickness have been obtained on Pt electrodes, which is ~ 100 times less than that typically available from lead zirconate-titanate with ruthenium oxide electrodes^{14–18}. (This large leakage current probably prevents the use of PZT in DRAMs but may not be a problem for non-volatile applications.) The switched charge density in most of the Sr/Bi layered-structure materials with optimum processing can be $25\text{--}32 \text{ } \mu\text{C cm}^{-2}$, which is more than adequate for non-volatile RAM device application^{19,20}. The retention properties of the new materials are significantly better than those of other ferroelectric materials measured at the film thicknesses and temperatures required for commercial devices^{19,20}. Typical properties are shown in Fig. 3, where the voltages produced in a load resistor are compared: the lower trace is the linear response of a ferroelectric already in the +P state to a positive voltage (no switching), the upper trace is the nonlinear (switching) response of a ferroelectric in the -P state to a positive voltage. 'Retention' is lost if the memory cannot discriminate between the two signals.

The absence of lead is an asset in these materials. Not only does this eliminate concerns about lead toxicity²⁶—for example, waste disposal problems and concerns for the safety of workers—it avoids the problem of lead vacancies, lead diffusion and Pb^{3+} states. Such defects can pair with neighbouring oxygen vacancies and other point defects to create orientable defect dipoles that also contribute to fatigue²⁷ in compounds such as PZT. In the materials under discussion, these sources of fatigue are also eliminated or minimized (Sr and Ba are much less volatile than Pb; and they are much less prone to exist in valence states other than divalent). Figure 2 illustrates the lack of fatigue out to 10^{12} cycles at 280 kV cm^{-1} . PZT will also operate without fatigue to $>10^{12}$ cycles, but at the cost of increased leakage current²⁸. It is not always possible to compare fatigue data from different laboratories; generally, neither thick films nor conducting films exhibit fatigue^{29,30}.

In addition to good retention and lack of fatigue, the importance of these new materials is the ability to fabricate very thin devices ($<100 \text{ nm}$) which have the same remanent polarization characteristics as thicker (or even bulk) materials. This may involve alleviation of mechanical stresses and strains, as well as injected-charge problems: bismuth oxide is well known to minimize interlayer stress^{31,32} in other layered compounds (including high- T_c superconductors), and it is also known that intergrowth superlattice compounds have a unique ability to convert point defects into strain lattice planes³³, which are less likely to pin domains and degrade switched charge^{2–4}. 'Tuning' the structure of these compounds can also be used to optimize breakdown voltage in commercial memory devices. We find both $2/1/2/1$ and $2/2/2/2$ stable intergrowth lattices in $\text{SrBi}_2\text{NbTaO}_9$, according to deposition conditions ($2/2$ designates two Bi_2O_3 planes followed by two TaO_6 octahedra). The $2/1/2/1/2/1$ superlattice exhibits higher breakdown voltages ($130 \pm 7 \text{ MV m}^{-1}$ at a thickness of 95 nm), essential in DRAMs intended to operate with 3 V across 30 nm (100 MV m^{-1}). □

Received 28 July 1994; accepted 9 March 1995.

1. Bondurant, D. W. & Gnadinger, F. P. *IEEE Spectrum* **26**, 30–34 (1989).
2. Artt, G. & Pertsev, N. A. *J. appl. Phys.* **70**, 2283–2289 (1991).
3. Artt, G. & Robels, U. *Integ. Ferroelect.* **3**, 247–254 (1993).
4. Plessner, K. W. *Proc. phys. Soc. B* **69**, 1261–1269 (1956).
5. Scott, J. F. & Araujo, C. A. *Science* **246**, 1400–1405 (1989).
6. Duiker, H. M. et al. *J. appl. Phys.* **68**, 5783–5789 (1990).
7. Postnikov, V. S., Pavlov, V. S., Gridnev, S. A. & Turkov, S. K. *Sov. Phys. Solid St.* **10**, 1267–1270 (1968).
8. Kudzin, A. Yu., Panchenko, T. V. & Yudin, S. P. *Sov. Phys. Solid St.* **16**, 1589–1591 (1975).
9. Stewart, W. C. & Cosentino, L. S. *Ferroelectrics* **1**, 149–153 (1970).
10. Lohkampfer, R., Neumann, H. & Artt, G. *J. appl. Phys.* **68**, 4220–4227 (1990).
11. Dakin, T. W. & Berg, D. *Progress in Dielectrics* Vol. 4 (eds Birks, J. B. & Hart, J.) 165–180 (Academic, New York, 1962).
12. Robblee, L. S. & Cogan, S. F. in *Encyclopedia of Materials Science and Engineering Suppl.* Vol. 1 (ed. Cahn, R. W.) 276–281 (Pergamon, Oxford, 1988).
13. Robblee, L. S. US Patent Nos 4677989 & 4717581 (1987).
14. Takasu, M. *Integ. Ferroelect.* (in the press).
15. Kwok, C. K., Vijay, D. P., Desu, S. B., Parikh, N. R. & Hill, E. A. *Integ. Ferroelect.* **3**, 121–128 (1993).
16. Krusin-Elbaum, L. & Wittmer, M. J. *electrochem. Soc.* **135**, 2610–2618 (1988).
17. Hren, P. D. et al. *Integ. Ferroelect.* **2**, 311–326 (1992).
18. Kwok, C. K., Vijay, D. P., Desu, S. B., Parikh, N. R. & Hill, E. A. *Integ. Ferroelect.* **3**, 121–130 (1993).
19. Fazan, P. C. *Integ. Ferroelect.* **4**, 247–256 (1994).
20. Kinney, W. *Integ. Ferroelect.* **4**, 131–144 (1994).
21. Smolenskii, G. A. & Agronovskaya, A. I. *Fiz. Tverd. Tela* **1**, 169–172, 442–445, 990–994 (1959); *Sov. Phys. Solid St.* **1**, 149–152 (1959).
22. Smolenskii, G. A. & Agronovskaya, A. I. *Fiz. Tverd. Tela* **1**, 442–445 (1959); *Sov. Phys. Solid St.* **1**, 400–403 (1959).
23. Smolenskii, G. A. & Agronovskaya, A. I. *Fiz. Tverd. Tela* **1**, 990–994 (1959); *Sov. Phys. Solid St.* **1**, 907–911 (1959).
24. Mihara, T. *Nikkei Electron.* **581**, 94–97 (1993).
25. Sumi, T. et al. in *Int. Solid St. Circuits Conf. (ISSC94)* (IEEE, New York, in the press).
26. Van Buskirk, P. C., Roeder, J., Bilodeau, S., Pombrink, S. & Beratan, H. *Integ. Ferroelect.* **6**, 141–153 (1995).
27. Smyth, D. in *Proc. NATO Conf. Science and Technology of Thin Films* (eds Auciello, O. & Waser, R.) (Kluwer Academic, Dordrecht, in the press).
28. Abt, N. E., Mistic, P., Zehngut, D. & Regan, E. *Integ. Ferroelect.* **4**, 349–355 (1994).
29. Abt, N. E., Moazzami, R. & Nissan-Cohen, Y. *Integ. Ferroelect.* **2**, 121–131 (1992).
30. Bernacki, S. et al. *Integ. Ferroelect.* **3**, 1–16 (1993).
31. Toledano, J.-C., Schneck, J. & Pierre, L. in *Incommensurate Modulations in Bismuth-Based High- T_c Superconductors* 335–345 (ed. Toledano, J.-C.) (NATO ASI Ser. B229, Plenum, New York, 1990).
32. Le Page, Y., McKinnon, W. R., Tarascon, J. M. & Barboux, P. *Phys. Rev.* **B40**, 6810–6816 (1989).
33. Anderson, J. S. & Hutchinson, J. L. *Commun. Phys.* **10**, 443–467 (1975).

Resolution beyond the 'information limit' in transmission electron microscopy

P. D. Nellist, B. C. McCallum
& J. M. Rodenburg

Cavendish Laboratory, University of Cambridge, Madingley Road,
Cambridge CB3 0HE, UK

THE conventional resolution of transmission electron microscopes is orders of magnitude larger than the wavelength of the electrons used. Aberrations of the objective lens corrupt spatial information on length scales below a limit known as the point resolution. Methods to correct for lens aberrations^{1–5} require knowledge of the phase of the waves which make up the image (this constitutes the 'phase problem'). Beyond the point resolution, information can still be transferred by the microscope, but partial coherence of the scattered beams imposes an ultimate limit (the 'information limit') on the resolution of the transferred image information. Here we show that this limit can be overcome to obtain images of still higher resolution with a scanning transmission electron microscope. Our approach involves collecting coherent microdiffraction patterns as a function of probe position, enabling us to extract the phase differences of all neighbouring pairs of diffracted beams. Using this approach for a microscope with a conventional point resolution of 0.42 nm and a conventional information limit of 0.33 nm, we are able to form an aberration-free image that resolves an atomic spacing of 0.136 nm.

Many solutions to the phase problem in electron microscopy exist. Indeed, holography was originally invented for this purpose^{1,2}. However, techniques such as off-axis holography^{3,4} and focal series reconstructions⁵ are still limited by the coherence of the microscope. Conditions of partial coherence are caused by the finite size of the electron source, the energy spread of the incident electrons, and other mechanical or electrical instabilities. These conditions limit the angle between diffracted beams which may interfere to form the image, thus creating an effective aperture which defines the 'information limit'⁶. Conventionally, no information beyond this effective aperture is transferred by the microscope, whether it be aberrated or not. The method used here is not limited by the information limit because it only requires interference to occur between adjacent diffracted beams which need not be very different in scattering angle.

In essence, our new approach solves the phase problem by restricting the area of specimen illuminated. This idea was first proposed by Hoppe^{7–9} and has been given the name 'ptychography'¹⁰ but it has never been successfully used to reconstruct high-resolution image information. The method of Konert *et al.*¹¹ employs a similar data set to the one used here, though they removed the central beam from the data. They proceeded by matching experimental data with calculations from a model structure, whereas here we perform direct phase determination with no *a priori* information on the structure under observation.

Crystalline silicon oriented on the $\langle 110 \rangle$ zone axis is a useful resolution test for transmission electron microscopes^{12,13}. In this projection the smallest apparent interatomic distance is 0.136 nm, which is at the limit of resolution of most state-of-the-art microscopes. Conventional images are affected both by the effects of multiple scattering of electrons within the specimen and by the lens aberrations^{14,15}. Here we use a scanning transmission electron microscope (STEM) to focus a demagnified image of a bright field-emission electron source onto an ~ 4.5 nm-thick crystal of silicon (Fig. 1). In the specimen plane, the focused

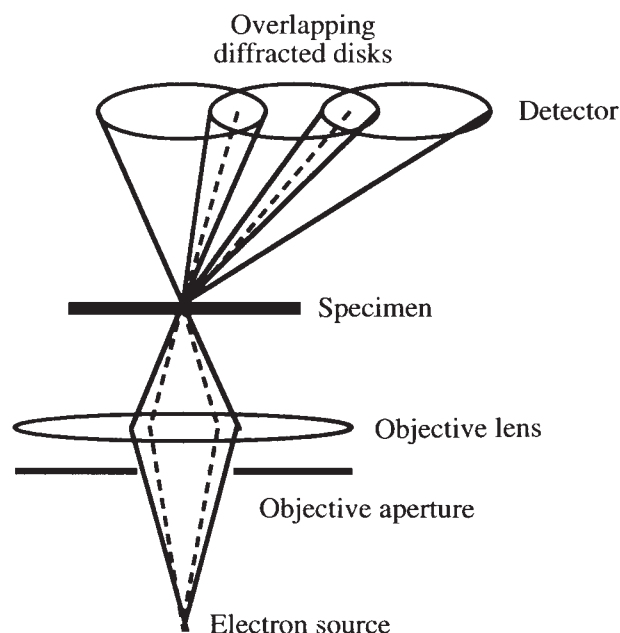


FIG. 1 Schematic diagram of the principle of aperture synthesis in STEM. By reciprocity^{25,26}, the observation of the periodic intensity fluctuation in the overlap as a function of probe position is equivalent to observing two-beam lattice fringes in a conventional transmission electron microscope with illumination tilt. The spatial frequency of such fringes is dependent only on the angle between the diffracted beams, not on their absolute scattering angle. For symmetric lens aberrations, the determined phase is aberration-free because the two waves which interfere at the centre of an overlap (indicated by dashed lines) have passed through diametrically opposite points of the objective lens. Asymmetric lens aberrations have been shown²⁰ to cause only a translation of the reconstructed image.

beam of electrons is ~ 0.5 nm in diameter. For a sufficiently large objective aperture, the diffracted disks in the far-field beyond the specimen overlap and coherent interference occurs (Fig. 2a). As the probe is scanned across the specimen, the intensity at each point of the overlap region varies sinusoidally (Fig. 3), which is simply STEM lattice imaging¹⁶. The phase of each intensity variation is determined by the relative phase of the two interfering beams^{17–19}. We proceed by taking a Fourier transform of the data with respect to the probe position, and observe that significant magnitude only occurs at the three different spatial frequencies which correspond to the three relative reciprocal lattice vectors of this crystal projection. From the Fourier-transformed data we can extract the phase difference of all neighbouring pairs of diffracted beams. Hence the phases of all the recorded diffracted beams can be deduced. The phases are combined with magnitudes measured from the centres of the disks, and an inverse Fourier transform is then taken to form a real-space image (Fig. 2b).

It is important to note that the spatial frequency of the intensity variation in an overlap region is equal to the reciprocal lattice vector joining the two diffracted beams. Therefore the frequency of the intensity variation is not dependent on the absolute scattering angle of the beams (Fig. 3), and so the conventional resolution limits do not apply. Using an optical-bench analogue of the STEM, the technique has been shown to be robust to many instrumental imperfections²⁰. However, the coherence and probe stability must be sufficient for interference to be observed. The highest angle between beams which we require to interfere corresponds to a relative reciprocal lattice vector of (002). In real-space this corresponds to a lattice spacing

Received 21 December 1994; accepted 14 March 1995.

1. Gabor, D. *Nature* **161**, 777–778 (1948).
2. Gabor, D. *Proc. R. Soc. A* **197**, 454–487 (1949).
3. Lichte, H. *Adv. opt. Electron Microsc.* **12**, 25–91 (1991).
4. Orchowski, A., Rau, W. D. & Lichte, H. *Phys. Rev. Lett.* **74**, 399–401 (1995).
5. Van Dyck, D., Op de Beeck, M. & Coene, W. *Optik* **93**, 103–107 (1993).
6. Spence, J. C. H. *Experimental High-Resolution Electron Microscopy* (Oxford Univ. Press, New York, 1988).
7. Hoppe, W. *Acta crystallogr.* **A25**, 495–501 (1969).
8. Hoppe, W. & Strube, G. *Acta crystallogr.* **A25**, 502–507 (1969).
9. Hoppe, W. *Acta crystallogr.* **A25**, 508–514 (1969).
10. Hegerl, R. & Hoppe, W. *Ber. Bunsenges. phys. Chem.* **74**, 1148–1154 (1979).
11. Konner, J., D'Antonio, P., Cowley, J. M., Higgs, A. & Ou, H. J. *Ultramicroscopy* **30**, 371–384 (1989).
12. Izui, K., Furuno, S. & Otsu, H. *J. Electron Microsc.* **26**, 129–132 (1977).
13. Izui, K., Furuno, S., Nishida, T., Otsu, H. & Kuwabara, S. *J. Electron Microsc.* **27**, 171–179 (1978).

14. Glaisher, R. W., Spargo, A. E. C. & Smith, D. J. *Ultramicroscopy* **27**, 19–34 (1989).
15. Glaisher, R. W., Spargo, A. E. C. & Smith, D. J. *Ultramicroscopy* **27**, 35–52 (1989).
16. Spence, J. C. H. & Cowley, J. M. *Optik* **50**, 129–142 (1978).
17. Nathan, R. in *Digital Processing of Biomedical Images* (eds Preston, K. & Onoe, M.) 75–88 (Plenum, New York, 1976).
18. Spence, J. C. H. *Optik* **49**, 117–120 (1977).
19. Spence, J. C. H. *Scann. Electron Microsc.* **1**, 61–68 (1978).
20. McCallum, B. C. & Rodenburg, J. M. *Ultramicroscopy* **52**, 85–99 (1993).
21. Rodenburg, J. M. & Bates, R. H. T. *Phil. Trans. R. Soc. A* **339**, 521–553 (1992).
22. McCallum, B. C. & Rodenburg, J. M. *Ultramicroscopy* **48**, 371–380 (1992).
23. Rodenburg, J. M., McCallum, B. C. & Nellist, P. D. *Ultramicroscopy* **48**, 303–314 (1993).
24. McCallum, B. C. & Rodenburg, J. M. *J. opt. Soc. Am A* **10**, 231–239 (1993).
25. Cowley, J. M. *Appl. Phys. Lett.* **15**, 58–59 (1969).
26. Cowley, J. M. *Ultramicroscopy* **2**, 3–16 (1976).
27. Dowell, W. C. T. & Goodman, P. *Phil. Mag.* **28**, 471–473 (1973).
28. Vine, W. J., Vincent, R., Spellward, P. & Steeds, J. W. *Ultramicroscopy* **41**, 423–428 (1992).

ACKNOWLEDGEMENTS. This work was supported by the UK EPSRC (formerly SERC), The Royal Society, St. John's College Cambridge, and the Cambridge Philosophical Society.

Geomorphological signatures of varying climate

Andrea Rinaldo*, William E. Dietrich†, Riccardo Rigon‡§, Gregor K. Vogel§ & Ignacio Rodriguez-Iturbe§

* Istituto di Idraulica "G. Poleni", Università di Padova, via Loredan 20, I-35131 Padova, Italy

† Department of Geology and Geophysics, University of California, Berkeley, California 94720, USA

‡ Dipartimento di Ingegneria Civile e Ambientale, Università di Trento, Mesiano di Povo I-38050, Italy

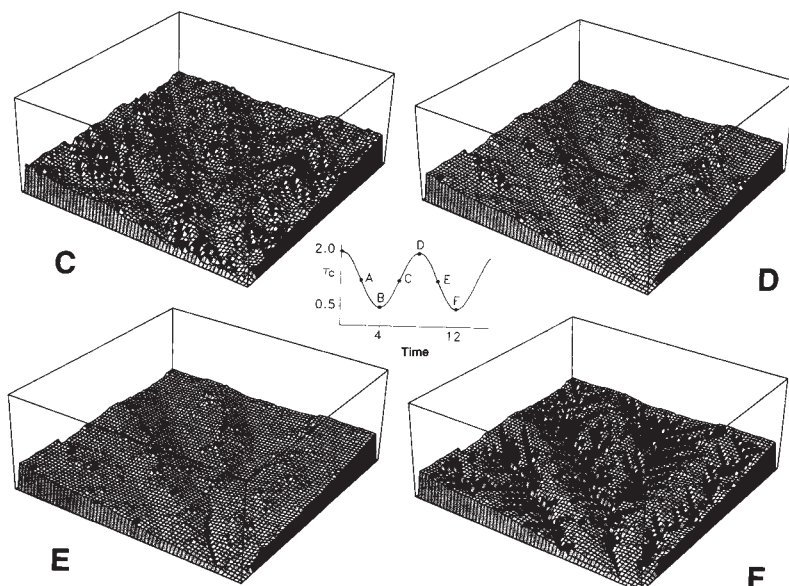
§ Department of Civil Engineering, Texas A&M University, College Station, Texas 77843, USA

A LONG-STANDING question in geomorphology^{1,2} is whether the topography of a particular landscape is in balance with current climate-driven processes, or contains relict signatures of past climates. For the glaciated landscapes of the Northern Hemisphere, the latter obviously applies but the situation is far from clear in regions where climate-driven processes have changed only in intensity, rather than character. We have addressed this question using a mathematical model of landscape evolution^{3,4} and find that both cases—contemporary balance and relict features—are pos-

sible. For a sinusoidal climate fluctuation, we find that all climate states (wet and dry) leave geomorphological signatures only when there is no active uplift. With active uplift and the associated increase in erosion, the topography tracks the current climate and any relict features are likely to reflect only the wettest conditions previously experienced by the landscape. In both cases, the temporal evolution of the landscape in response to cyclic climate forcing is complex, and leads to the unexpected result that valley density is largest during periods dominated by slow downslope movement of sediment, rather than during times of strong fluvial incision, as would be anticipated from steady-state models^{5–7}.

Our mathematical model of landscape self-organization^{3,4} treats the interplay of two erosion processes: relatively fast surface erosion by running water (fluvial processes) and slow downslope movement that depends on local surface gradient (diffusive transport). We treat the erosion by fluvial processes as occurring once a critical shear stress (τ_c) has been exceeded, that is, the net flux is proportional to a function $f(\tau - \tau_c)$ where $f(x) = 0$ for $x \leq 0$. The shear stress $\tau(X, t)$ at a site X at time t is assumed to depend on the total drainage area A (determined through steepest-descent drainage directions) and slope ∇z , where $z(X, t)$ is the landscape elevation field, via $\tau \propto \sqrt{A} \nabla z^3$. Area acts as a surrogate variable for mean annual discharge at a site. Thus shear stresses are objectively computed once the topographical field z is determined. All sediment mobilized by this process is removed from the system—no re-deposition can occur. We refer to this process as threshold-limited.

FIG. 1 An example of the effects of climate fluctuations on landscape self-organization. The outcome of the model is a topographic field $z(X, t)$ on a lattice, where z is the point elevation of the landscape, X is a cartesian coordinate and t is time. We simply alternate erosion cycles (in which all sites in excess of the critical shear stress (τ_c) are reduced to this critical value, and thus the system self-tunes from thousands of threshold-limited changes to very few) at fixed or random intervals with cycles of diffusive transport (in which a predetermined number of transport steps are chosen for entire lattice). In the model used here, 10,000 individual site slope-dependent calculations are employed, implying that in the average each pixel undergoes about three slope-dependent changes per diffusive cycle in a 64×64 lattice. Elsewhere⁸ we show that our results are independent of the particular choice of number of diffusive events per cycle. The threshold-limited landscape is taken as the initial condition for the next slope-dependent diffusive evolution and vice versa. The process can be iterated arbitrarily. Here we do not attempt to relate the duration of diffusion between instantaneous threshold-limited erosion to specific magnitude or frequency scales of climate oscillations. Our conclusions are seen as independent of the particular scales as well as of the particular choice of initial conditions³. For four selected stages of evolution we show (for the case $U=0$) the resulting elevation fields, dated with letters from C to F with reference to the current value of τ_c (shown



in the inset) used to update threshold-limited erosion at regular time intervals.

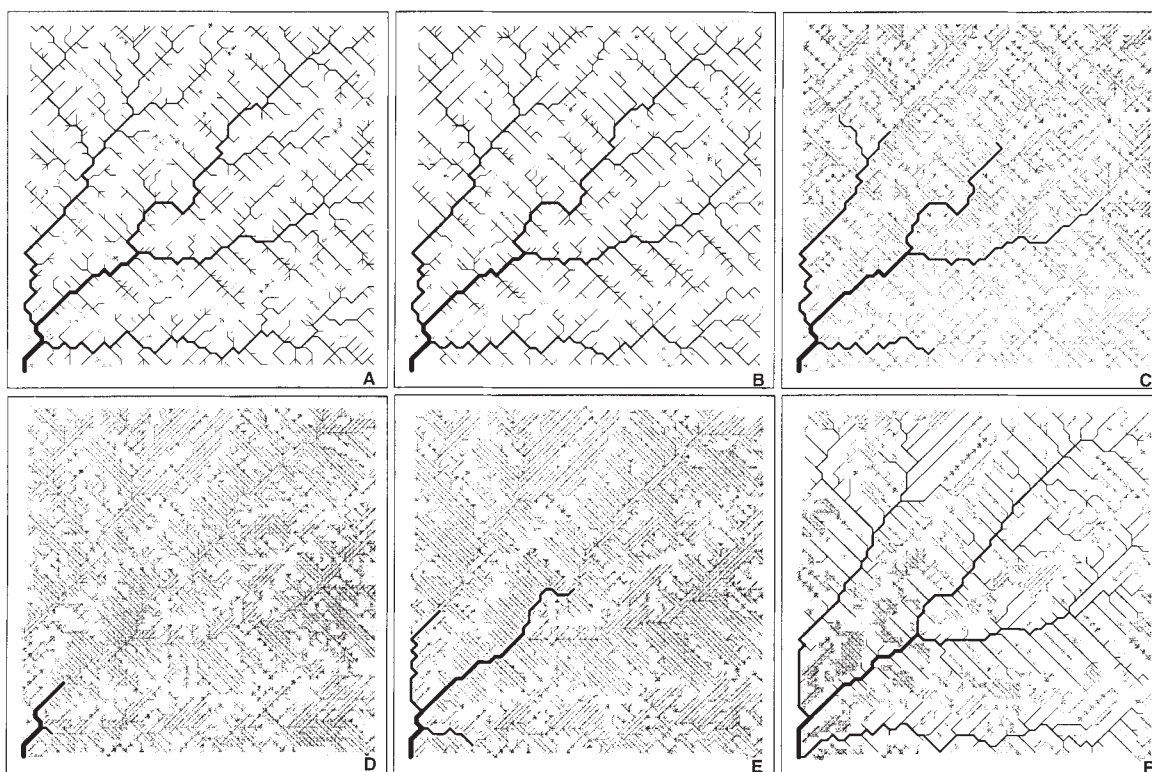


FIG. 2 Drainage directions on channelled and unchannelled valley sites corresponding to the landscapes A to F with reference to the current value of τ_c (see Fig. 1, inset). Drainage directions are defined through the steepest descent ∇z defined in the lattice³. Here shaded pixels indicate valleys as areas of convergent topography, defined²², at time

t , by the condition $\nabla^2 z(X, t) \geq 0$. Channelled valleys (defined^{17,22,24} by $\nabla^2 z \geq 0$ and $\tau \geq \tau_c(t)$) are described by solid lines whose thickness is proportional to $\sqrt{A}$. Dotted lines indicate drainage directions in unchannelled regions. Convex hillslopes are represented by blank pixels.

Diffusive sediment transport depends on local surface slope (∇z) and this slope-dependent process affects the rate of change in z in conjunction with geological uplift (U). Mobilized sediments can be re-deposited by this process. The characteristic timescales of threshold-limited fluvial erosion are assumed to be much smaller than those of slope-dependent transport, and thus we may decouple⁴ the instantaneous threshold-limited erosion from the slope-dependent modification of landforms. Despite the simplicity of the model, the competition between threshold-limited valley incision and slope-dependent smoothing of ridges has successfully reproduced^{3,4,8} the complex patterns of spatial self-organization of large interactive dynamical systems⁹⁻¹⁴, as well as many properties of real landscapes¹⁵⁻²⁴.

We propose to simulate climate fluctuations by imposing a cycle on the threshold value $\tau_c = \tau_c(t)$. Of course, climate oscillations are not simple cycles, and landscape response to (for example) varying precipitation can have significant lags and complexities associated with vegetation response. We suggest, however, that the very simplicity of this model enables us to identify causal relationships in landscape response and to capture the essence of the combined effects of climate changes. As discussed later, the complexity of the cycle structure does not alter substantially our main findings. Here we interpret low τ_c and high τ_c as representing wetter and drier conditions, respectively. We assume that changes in slope-dependent transport due to climate changes are negligible compared to those occurring in rates of threshold-limited transport owing to varying critical shear stresses.

Figures 1 and 2 show a complete example of the effects of a sinusoidal climate fluctuation where a landscape undergoes a major change in climate reflected by a large variation in the critical shear stress. In Figs 1 and 2 the uplift rate U is set to zero. We observe that landscape states corresponding to equal thresholds (say A and E, and B and F, as defined on the sinu-

soidal graph in Fig. 1) differ significantly as the system evolves, because the progressively declining mean slope causes the critical shear stress to be less frequently exceeded. Unchannelled valleys (Fig. 2) are created by this oscillation in a manner suggested by field observations²³. We have repeated the cycle shown in Fig. 1 for two different uplift rates U , here defined in pixel units as respectively 0.4 and 0.8 height units per climate cycle. Other simulations based on more complex fluctuations built on different climate frequencies, which we will report elsewhere, confirm our present conclusions.

A summary of our results is given in Fig. 3. We use three geomorphological indicators to describe the effects of climate oscillations on landscape morphology: a global fractal dimension²⁵ of the elevation field; the drainage density¹⁵ of the resulting basins (the total length of channels divided by landscape area—see Fig. 3 legend for details); and the related valley density²². In this context, valley density measures the proportion of pixels with concave topography ($\nabla^2 z \geq 0$). In general, during rising values of critical shear stress, hillslope (slope-dependent) processes tend to smooth the landscape, and during falling values threshold-limited fluvial processes may roughen it. In the absence of uplift ($U=0$) the roughness of the landscape (and thus its fractal dimension) shows a damped oscillation which is not observed when uplift occurs (Fig. 3b). Without uplift, progressive flattening of the landscape causes threshold-limited processes to become less active while the diffusion-like transport continues; this prevents restoration of the higher fractal dimensions during periods of low critical stress. During rising critical shear-stress values (equivalent to aridification in this model), slope-dependent processes tend to infill the previous channel networks, creating an extensive network of unchannelled valleys (where shear stress is below critical), but not altering substantially the overall valley topography created during the previous low critical shear-stress (humid) cycle. For both the case of 'no

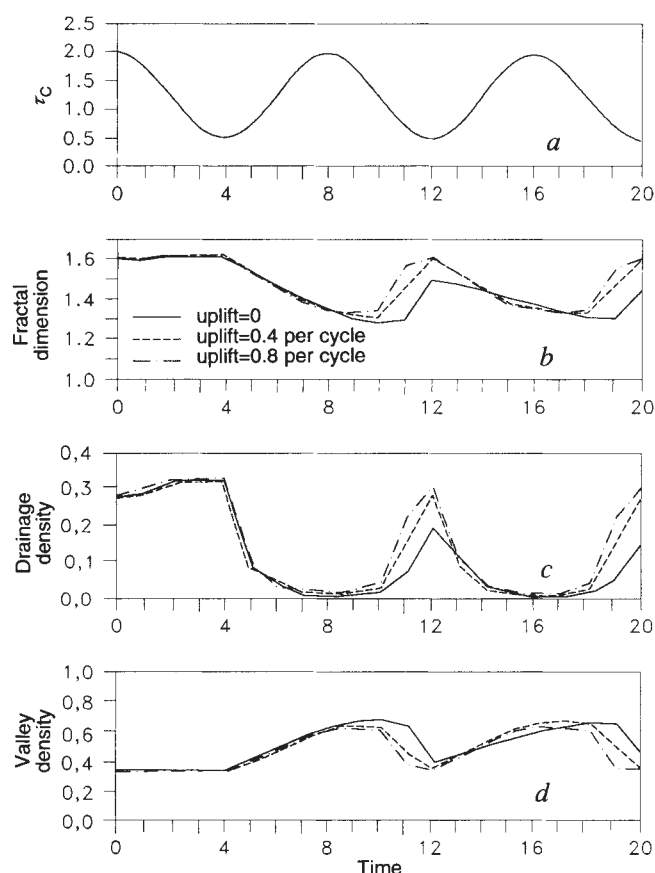


FIG. 3 For three different uplift rates ($U=0, 0.4$ and 0.8 height units per cycle), time evolution of: a, the critical shear stress $\tau_c(t)$ defining the climate cycle; b, the fractal dimension²⁵ of the elevation field, which measures the roughness of the landscape and is determined by generalized covariance analysis⁸. A single fractal dimension postulates a simple scaling^{25–27} elevation field and although natural topographies are self-organized landscapes are generally multifractal^{8,28}, the simple scaling assumption proves viable and robust⁸; c, the drainage density which is a measure of the degree to which the basin is dissected by channels. It is usually defined¹⁵ by the ratio of the total length of channels to the total area of the basin. Here we employ a dimensionless ratio defined as the ratio of the number of pixels (area units making up the lattice used in the simulations) where a channel occurs divided by the total number of pixels. Experimental^{17,20,24} and theoretical^{4,17,22,23} evidence strongly suggests that a site i is occupied by channels if $\tau_i = \sqrt{A_i} \nabla z_i \geq \tau_c$ and simultaneously the site is located in a concave valley region where $\nabla^2 z \geq 0$; d, the valley density of the basin, defined²² as the ratio of the number of concave sites (that is, where $\nabla^2 z \geq 0$) defining valley regions of convergent topography (usually identified with colluvium regions) and the total number of sites in the basin. Time here is given in arbitrary units defined by the number of slope-dependent changes employed (10,000 slope-dependent changes correspond to one-quarter of the climate cycle, that is, two time units).

uplift' or 'active uplift', the valley density (Fig. 3d) is less variable than the drainage density. In this model, then, it is the timing and magnitude of drainage-density change in response to climate oscillations that leads to landform inheritance. Figure 3c shows that drainage density increases (more of the basin becomes channelized) as the climate becomes more humid (lower τ_c), and decreases as the climate becomes more arid (higher τ_c) when the channels nearly disappear. With active uplift, when the critical shear stress is periodically lowered to similar values in successive cycles, all past drainage paths are recut. Without tectonic uplift ($U=0$), however, drainage density during successive wet periods progressively declines.

Important results emerge. During periods of weakened river incision, but still with active diffusive processes, the stable base

levels of the valley bottoms cause the concave area to expand by diffusive processes. Although it is well established^{7,23} that concavity results from threshold-limited processes (and convexity originates from diffusive slope-dependent processes), when the climate forcing is periodic, as will occur virtually anywhere, we actually see periods of greater concavity associated with periods of weakened river incision. Importantly, the expansion of the concave region is not due to hillslope sheet wash, but rather to diffusive processes.

The response function of valley density to cyclic forcing (Fig. 3d) shows an intriguing asymmetry reflected in the short lag occurring between the maximum and the minimum values. If this is similar to what happens in nature, it suggests how significant time lags between climate change and landform adjustment might occur. With the cyclic re-establishment of conditions favourable to threshold-limited erosion after the peak of the dry phase (the descending limb of $\tau_c(t)$, Fig. 3a), valleys again become active with channels that advance headwards and cut down. During this advancement phase, diffusive processes keep wearing the slopes and expanding the area of concave topography, effectively increasing the valley area. The valley density decreases only when the channels extend up to their former maximum extension and begin cutting down, because only then the convex hillslopes associated with basal incision expand downslope. This explains the unforeseen lag and the short time incurred between maximum and minimum valley density. Thus the extent of convexity of the topography is positively correlated with channel incision, rather than negatively as one would have otherwise expected from the viewpoint of competition of processes under steady-state conditions. This means that a landscape which is on the average in steady state (that is, the average uplift and erosion are equal), but which is experiencing periodic climate forcing, may be radically different in morphology than one which is strictly in steady state (no variation in uplift or climate forcing). Morphological differences are especially manifest in the extent of unchannelled valleys, because valley and drainage densities are sometimes both increasing and at other times changing in opposite directions. Thus channel initiation patterns not only set a finite scale to the landscape^{17,20} (and a lower cutoff^{4,17} to the growth of fluvial fractal structures^{3,25}) but also play a complex role in the interpretation of the geomorphological signatures of past climates.

The change of drainage density with the climate variable $\tau_c(t)$ in the case $U=0$ clearly shows hysteresis, suggesting that the entire climate pattern affects the current geomorphology. With active tectonics, however, once the initial conditions are erased the particular drainage density is purely a function of the specific climate. The extent of unchannelled valleys reflects just the previous threshold minimum, and does not tend to increase in extent through successive cycles.

We suggest that, given the apparent lack of preferential scale in the climate record (that is, a power-law behaviour of climate fluctuation rather than a smooth sinusoidal pattern), extreme climate excursions that cause accelerated threshold-limited erosion seem likely to occur and to leave very long-lived morphological evidence where slope-dependent processes are ineffective. This may explain such features as rock pediment growth and the persistence of ancient fluvial landforms in hyper-arid climates. □

Received 1 December 1994; accepted 9 March 1995.

1. Summerfield, M. A. *Global Geomorphology* Ch. 14 (Longman, Singapore, 1991).
2. Bloom, A. L. *Geomorphology: a Systematic Analysis of Late Cenozoic Forms* (Prentice Hall, New Jersey, 1991).
3. Rinaldo, A., Rodriguez Iturbe, I., Rigon, R., Ijjasz Vasquez, E. & Bras, R. L. *Phys. Rev. Lett.* **70**, 822–826 (1993).
4. Rigon, R., Rinaldo, A. & Rodriguez Iturbe, I. *J. geophys. Res.* **99**, 11971–11993 (1994).
5. Kirkby, M. J. *Slopes: Form and Process* (Spec. Publ. No. 3, Inst. of British Geographers, London, 1971).
6. Willgoose, G. R., Bras, R. L. & Rodriguez-Iturbe, I. *Wat. Resour. Res.* **27**, 1671–1702 (1991).
7. Kirkby, M. J. in *Channel Network Hydrology* (eds Beven, K. & Kirkby, M. J.) 253–293 (Wiley, New York, 1993).

8. Rodríguez-Iturbe, I., Marani, M., Rigon, R. & Rinaldo, A. *Wat. Resour. Res.* **30**, 3531–3540 (1994).
9. Bak, P., Tang, C. & Wiesenfeld, K. *Phys. Rev. Lett.* **59**, 381–385 (1987).
10. Bak, P., Tang, C. & Wiesenfeld, K. *Phys. Rev. A* **38**, 364–374 (1988).
11. Bak, P., Chen, K. & Creutz, M. *Nature* **342**, 780–781 (1989).
12. Bak, P. & Paczuski, M. *Physics World* **6**, 39–43 (1993).
13. Rodríguez-Iturbe, I., Rinaldo, A., Rigon, R., Bras, R. L. & Ijász-Vasquez, E. *Wat. Resour. Res.* **28**, 1095–1103 (1992).
14. Rinaldo, A. et al. *Wat. Resour. Res.* **28**, 2183–2195 (1992).
15. Horton, R. E. *Bull. geol. Soc. Am.* **56**, 275–370 (1945).
16. Tarboton, D. G., Bras, R. L. & Rodríguez-Iturbe, I. *Wat. Resour. Res.* **24**, 1317–1322 (1988).
17. Montgomery, D. R. & Dietrich, W. E. *Nature* **336**, 232–234 (1988).
18. Tarboton, D. G., Bras, R. L. & Rodríguez-Iturbe, I. *Wat. Resour. Res.* **25**, 2037–2051 (1989).
19. Rodríguez-Iturbe, I., Ijász-Vasquez, E., Bras, R. L. & Tarboton, D. G. *Wat. Resour. Res.* **28**, 988–993 (1992).
20. Montgomery, D. R. & Dietrich, W. E. *Science* **255**, 826–830 (1992).
21. Uchii, S. & Matsushita, M. *Geomorphology* **5**, 115–130 (1992).
22. Dietrich, W. E., Wilson, C. J., Montgomery, D. R. & McKean, J. J. *J. Geol.* **3**, 161–180 (1993).
23. Dietrich, W. E. & Dunne, T. in *Channel Networks Hydrology* (eds Beven, K. & Kirkby, M. J.) 175–219 (Wiley, Chichester, 1993).
24. Montgomery, D. R. & Foufoula-Georgiou, E. *Wat. Resour. Res.* **29**, 3925–3934 (1993).
25. Mandelbrot, B. B. *The Fractal Geometry of Nature* (Freeman, San Francisco, 1983).
26. Lovejoy, S., Schertzer, D. & Ladoy, P. *Nature* **319**, 43–44 (1986).
27. Lovejoy, S. & Schertzer, D. *J. geophys. Res.* **95**, 2021–2034 (1990).
28. Lavalley, D., Lovejoy, S., Schertzer, D. & Ladoy, P. in *Fractals in Geophysics* (eds Lam, N. & De Cola, L.) 158–192 (Prentice-Hall, Englewood Cliffs, 1993).

Hepatic characters in the earliest land plants

D. Edwards*, J. G. Duckett† & J. B. Richardson‡

* Department of Earth Sciences, University of Wales College of Cardiff, PO Box 914, Cardiff CF1 3YE, UK

† School of Biological Sciences, Queen Mary and Westfield College, Mile End Road, University of London, London E1 4NS, UK

‡ Department of Palaeontology, Natural History Museum, Cromwell Road, London SW7 5BD, UK

EVIDENCE for terrestrial vegetation in Ordovician and Silurian times, before the advent of vascular plants, comes from palynomorphs (cryptospores¹), that is, non-dissociating tetrads and dyads^{2,3} and cuticles⁴. The lack of a megafossil record for the spore producers is usually attributed to low fossilization potential of vegetative tissues, and this, plus spore type⁵, contributes to the hypothesis that the plants were embryophytes/archegoniates at a bryophyte level of organization^{3,5,6} and perhaps most similar in organization to modern hepatics⁵. Here we describe a minute coalified fossil from the Lower Devonian (Lochkovian: *micronatus-newportensis* Spore Biozone) of the Welsh Borderland⁷, which contains obligate, smooth-walled tetrahedral tetrads, similar (by scanning electron microscopy) to those first recorded in the Ordovician. To our knowledge, these are the first data on the gross morphology and tissues of the plants that comprised the earliest embryophyte land flora (Gray's Eoembryophytic epoch³), albeit obtained from a relict Devonian example fossilized at a time when the composition of dispersed spore and megafossil assemblages suggests that tracheophytes and tracheophyte-like plants (rhyniophytoids) had generally begun to dominate land vegetation (Gray's Eotracheophyta³). Its anatomy *in toto* finds no exact parallels in embryophytes, but many of the individual cellular features match those in extant hepatics (liverworts).

The tetrads occur in two interconnecting lobes of a bifurcation at one end of a small coalified axial fossil about 1.54 mm long (Fig. 1a). Each crassitate unit (~30 µm diameter) with a depressed distal surface is united to adjoining ones by a narrow continuous layer which is invaginated to varying extents (Fig. 1b). From scanning electron micrographs (SEMs) it appears as if the tetrad is enveloped by a smooth adhering layer, but fractured examples show no evidence of a separate layer (although diagenetic homogenization is possible). Attempts to isolate and clear specimens for light microscopy failed, because the tetrads remain attached to thicker, opaque surrounding tissue. However the superficial features conform to those of non-dissociating tetrahedral spore tetrads with closest similarities with *Tetrad-*

raletes (Ordovician (Ashgill cited³: Caradoc illustrated⁴)–Lower Devonian), and among extant plants, the permanent smooth tetrads of marchantialean hepatics⁸. A single strap-shaped structure occurring among the spores (Fig. 1c) is suggestive of hepaticalean elaters.

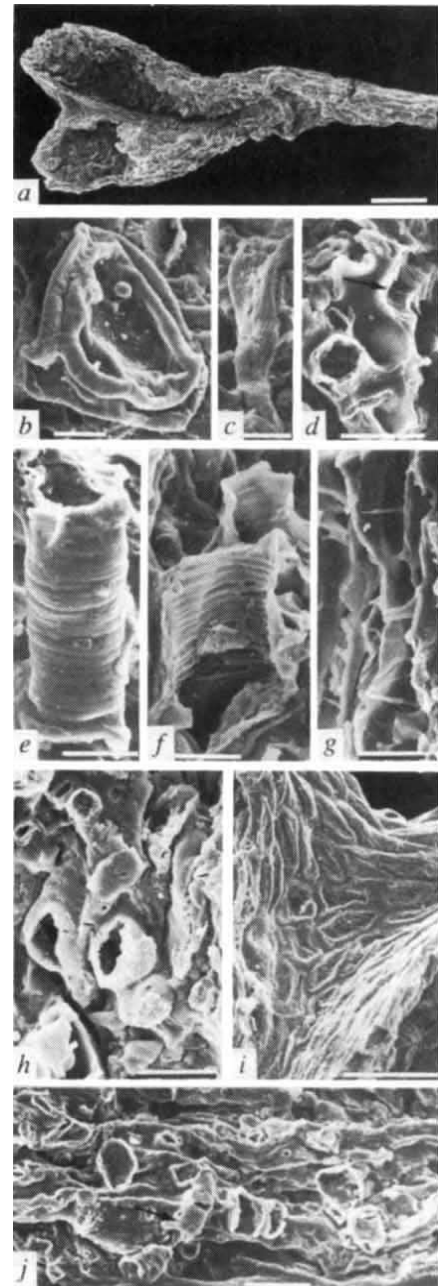


FIG. 1 a, SEM of complete coalified specimen isolated from a Lochkovian siltstone, Welsh Borderland using hydrofluoric acid. Arrow indicates position of cells in *i*. NMW 94. 76G. 1. b, *In situ* tetrad. c, Elater-like structure, with central groove and herring-bone wall ingrowths, located near spores. d, Fragment of internal tissue, with type 2 tube arrowed. e, Part of type 1 tube in vicinity of spores, ± transversely orientated. f, Part of internal face of type 2 tube, 'plunging' into tissue. cf, *Porcatitubulus*. g, Longitudinally aligned elements with irregular wall ingrowths (type 3); h, ? remains of spherical cells. i, Superficial cells with more clearly striated central areas. j, Typical appearance of 'ground tissue' with large cup shaped structures and possible pit fields. Inner surfaces of larger 'cups' more granular than outer and structures irregularly perforated. Superficial tube 1 arrowed. Scale bars: a, 200 µm; b–h, 10 µm; i, 50 µm; j, 20 µm.

The only readily recognized superficial cells occur in the angle between spore-containing areas and around their bases (Fig. 1i). These gradually merge proximally into the disorganized and apparently superficial axial tissues, lacking any evidence for a cuticle. This is in marked contrast with the cuticularized smooth or regularly ridged surfaces of tracheophytes and rhyniophytoids in the assemblage⁷. The most conspicuous superficial features are cup-shaped outgrowths with jagged margins sometimes incurving, suggestive of once-spherical structures (Fig. 1h, j), although also possibly the swollen bases of rhizoids⁹. However they are more numerous, being either scattered or clustered when present in deeply seated highly disorganized tissues. Well-defined smooth-walled cells or tubes have not been observed, but a number of elongate elements with unevenly thickened walls include the following. (1) Tubes characterized externally by narrow grooves delimiting broad annulations (Fig. 1e), which appear as low thickenings on the inside. They are usually circular in cross-section, rarely flattened and, apart from one adhering to the surface (Fig. 1j), are scattered throughout the axis and relatively isolated from the other tissues. (2) Tubes with irregular and narrower internal thickenings, more intimately related to the disorganized tissues (Fig. 1f). (3) Predominantly longitudinally aligned forms, of uncertain position in the fossil, with irregular wall ingrowths (strands and ledges) (Fig. 1g). Crater-shaped structures (2–4 µm wide) seen on fractured surfaces are interpreted as surface views of pit fields in these and possibly other tissues. (4) Peripheral thick-walled cells, with homogenized walls and small lumina marked by transversely orientated furrows (not illustrated). This peripheral zone is C-shaped in transverse section. Such incomplete preservation may be fortuitous or may reflect differences in original tissue composition, for example chlorenchyma versus structural tissue.

In considering the affinities of the fossil, its gross morphology resembles that of coeval tracheophytes and rhyniophytoids (for example, *Salopella marcensis*⁷, a plant with rhyniophyte morphology, but in which tracheids have not been demonstrated). Although spores and vegetative organization differ, unevenly thickened cells characterize rhyniophytoid peripheral structural tissues (as in *Tortilicaulis*⁷), whereas the more regularly thickened type 2 described above resembles annular/spiral tracheids except that the wall between thickenings is thicker and more resilient. In this respect, they are close to tracheids of *Cooksonia pertonii*¹⁰ and indicate how xylem might have evolved from putative structural elements by progressive loss of wall to accommodate extension growth and water exchange, with concomitant concentration into a vascular strand¹¹. The lack of well-defined tubular organization precludes nematophyte affinity (*sensu* Lang¹²), although unevenly thickened tubes occur in *Nematosketum*¹³ and isolated wefts¹⁴. Some of the banded tubes (for example, *Porcatitubulus* sp.) recovered from palynomorph assemblages from latest Llandovery¹⁵ onwards undoubtedly derive from plants with the new fossil's organization. It might also be related to a small sterile Llandovery axial fossil¹⁶.

Among the bryophytes, affinity with hepatics is supported by several anatomical features: (1) type-1 tubes may be interpreted as rhizoids, which today show banding within the wall⁹; (2) type-3 tubes recall irregularly thickened food-conducting cells in Marchantiales¹⁷; (3) the cup-shaped structures are reminiscent of scattered idioblastic oil-body cells in the same group; (4) permanent tetrads. These features encompass both gametophytic and sporophytic characters and lead to conflicting hypotheses on the nature of the fossil. Gross morphology is indeed suggestive of a sporophyte, but in extant hepatics this is unbranched and lacks tissue differentiation⁸. The other possibility is that the fossil represents a bifurcating thallus with an internalized sporophyte, as occurs in extant Ricciaceae, where only vestiges of the sporophyte remain at spore maturity. As in the fossil, rhizoids occur in the vicinity of the spores and grow into the decaying thallus tissue⁸.

In conclusion, this Lower Devonian fossil with novel organization, incorporating unevenly thickened tubes (first recorded in dispersed assemblages in the latest Llandovery) and permanent spore-tetrads with similar spores extending back into the Ordovician, provides a unique glimpse of a pre-tracheophyte level of organization. Current views on bryophyte phylogeny based on phenetic, cladistic^{18,19} and molecular²⁰ approaches underline the antiquity of the bryophytes, but are equivocal on the nature of the oldest members. It is not unexpected that the fossil described here has most similarities with thalloid hepatics but does not conform exactly to any subsequent group: such disparity might be anticipated in the group from which embryophytes radiated. Extant bryophytes are the products of over 400 million years of evolution, which is ample time for changes to have occurred, although with frustratingly poor documentation in the fossil record.

Note added in proof: Taylor²¹ has reported sphaerocarpacean-type ultrastructure in Lower Silurian dyads (*Dyadospore*), thus reinforcing arguments for a hepatic affinity for early land plants. □

Received 20 December 1994; accepted 21 February 1995.

- Richardson, J. B., Ford, J. H. & Parker, F. J. *Micropal.* **3**, 109–124 (1984).
- Gray, J. *Syst. Assoc. Spec.* **44**, 49–87 (1991).
- Gray, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **104**, 153–169 (1993).
- Gray, J., Massa, D. & Boucot, A. J. *Geology* **10**, 197–201 (1982).
- Gray, J. *Phil. Trans. R. Soc. B* **309**, 167–195 (1985).
- Gray, J. *Spec. Pap. Pal.* **32**, 281–294 (1984).
- Edwards, D., Fanning, U. & Richardson, J. B. *Bot. J. Linn. Soc.* (1994).
- Schuster, R. M. *New Manual of Bryology* Vol. 2 (Hattori Botanical Laboratory, Nichinan, Japan, 1984).
- Pocock, K. & Duckett, J. G. *New Phytol.* **99**, 281–304 (1985).
- Edwards, D., Davies, K. L. & Axe, L. *Nature* **357**, 683–685 (1992).
- Edwards, D. *New Phytol.* **125**, 225–247 (1993).
- Lang, W. H. *Phil. Trans. R. Soc. B* **227**, 245–291 (1937).
- Burgess, N. D. & Edwards, D. *Bot. J. Linn. Soc.* **97**, 189–203 (1988).
- Edwards, D. *Bot. J. Linn. Soc.* **84**, 223–256 (1982).
- Burgess, N. D. & Edwards, D. *Bot. J. Linn. Soc.* **106**, 41–66 (1991).
- Niklas, K. J. & Smocovitis, V. *Paleobiology* **9**, 126–137 (1983).
- Ligrone, R. & Duckett, J. G. *Ann. Bot.* **73**, 577–586 (1994).
- Garbary, D. J., Renzaglia, K. S. & Duckett, J. G. *Plant Syst. Evol.* **188**, 237–269 (1993).
- Mischler, B. D. & Churchill, S. P. *Cladistics* **1**, 305–328 (1985).
- Mischler, B. D., Thoall, P. H., Hoppie, J. S. Jr, De Luna, E. & Vilgalys, R. *Bryologist* **95**, 172–189 (1992).
- Taylor, W. A. *Nature* **373**, 391–392 (1995).

ACKNOWLEDGEMENTS. Research supported by NERC grants. We thank L. Axe and C. Wellman for technical assistance.

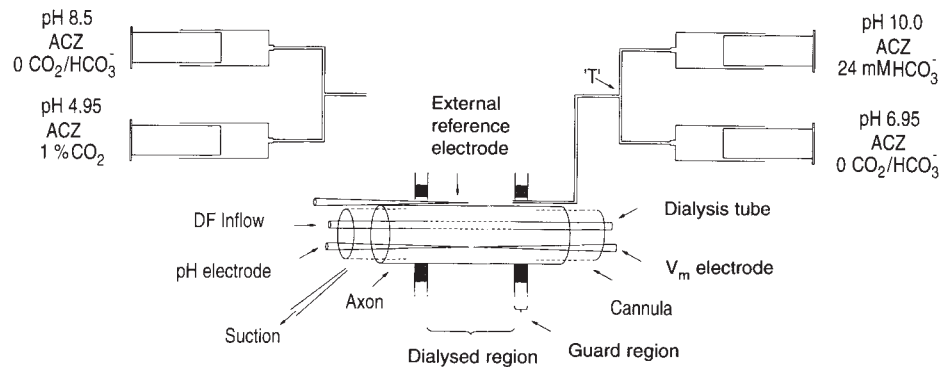
Out-of-equilibrium CO₂/HCO₃[−] solutions and their use in characterizing a new K/HCO₃ cotransporter

Jinhua Zhao, Emilia M. Hogan, Mark O. Bevensee & Walter F. Boron

Department of Cellular and Molecular Physiology,
Yale University School of Medicine, New Haven,
Connecticut 06510, USA

In typical physiological solutions, CO₂ is in equilibrium with HCO₃[−] and H⁺ (CO₂ + H₂O ⇌ HCO₃[−] + H⁺). Because one cannot independently alter CO₂ and HCO₃[−] concentrations and pH, it is impossible to distinguish between the effects of CO₂ and HCO₃[−] on physiological processes. Here we describe a continuous-flow, rapid-mixing approach for generating out-of-equilibrium CO₂/HCO₃[−] solutions with a physiological pH and CO₂ (but little HCO₃[−]), or pH and HCO₃[−] (but little CO₂). We have exploited these out-of-equilibrium solutions to introduce HCO₃[−] exclusively to either the outside or inside of a squid giant axon, and verify the

FIG. 1 Generation of out-of-equilibrium (OOE) solutions. Schematic diagram of continuous-flow apparatus and delivery of OOE solutions to a squid giant axon. The tubing downstream from mixing 'T' had a volume of 50 μl . The time from the T to the end of the delivery tube was ~ 475 ms (flow = 6.3 ml min^{-1}). When a 24 mM HCO_3^- solution at pH 10 ($[\text{CO}_2] = 2.3 \mu\text{M}$) is mixed with a $\text{CO}_2/\text{HCO}_3^-$ -free solution at pH 6.95, the computed $[\text{H}_2\text{CO}_3^*]$ at the T is $0.29 \mu\text{M}$, based on the pH (8.00), $[\text{HCO}_3^-]$ (12 mM) and equilibrium constant of the reaction $\text{HCO}_3^- + \text{H}^+ \rightleftharpoons \text{H}_2\text{CO}_3^*$ ($\text{pK} = 3.38$). The rate constant of the reaction $\text{H}_2\text{CO}_3^* \rightarrow \text{CO}_2 + \text{H}_2\text{O}$ is $\sim 15 \text{ s}^{-1}$ at 25°C (refs 19–21), so that CO_2 forms at a rate of $4.3 \mu\text{M s}^{-1}$. Thus, $2.1 \mu\text{M}$ of CO_2 forms after 475 ms, for a total $[\text{CO}_2]$ of $(2.3/2) + 2.1 = \sim 3.3 \mu\text{M}$. We neglected the back reaction because it is insignificant for the first 475 ms. As measured with a CO_2 electrode (model MI-720, Microelectrodes, Inc.; detection limit, 4–5 μM), $[\text{CO}_2]$ was ~ 0 at the end of the delivery tube. With perfect mixing in the chamber (volume = $\sim 5 \mu\text{l}$), the time constant for solution exchange would be ~ 50 ms, and only a small amount of additional CO_2 would form in the chamber. In fact, chamber $[\text{CO}_2]$ was sometimes much higher than $3.3 \mu\text{M}$ because of imperfect mixing. When a 1% CO_2 at pH 4.95 ($[\text{HCO}_3^-] = 240 \mu\text{M}$) is mixed with a $\text{CO}_2/\text{HCO}_3^-$ -free solution at pH 8.5, the $[\text{CO}_2]$ at the T is $\sim 112 \mu\text{M}$. Because the rate constant of $\text{HCO}_3^- + \text{H}^+ \rightarrow \text{H}_2\text{CO}_3^*$ is $\sim 0.0382 \text{ s}^{-1}$ at 25°C (refs 19–21), CO_2 is consumed at the rate of $4.3 \mu\text{M s}^{-1}$. Thus, $2.1 \mu\text{M}$ of HCO_3^-



forms after 475 ms, for a total $[\text{HCO}_3^-]$ of $(240/2) + 2.1 = \sim 122 \mu\text{M}$. METHODS. Generating OOE solution: two 140-ml plastic syringes were driven by a syringe pump (model 55-2222, Harvard Instruments). When generating an OOE HCO_3^- solution, we filled one syringe with a pH 10, 24 mM HCO_3^- solution, and the other with a pH 6.95/ HCO_3^- -free solution. The output of each syringe was directed through tygon tubing to an assembly of 5-way valves (model E4-1PP, Clippard), and then through a stainless-steel tube surrounded by a 22°C water jacket (not shown), and finally to the mixing T. Downstream from the T, solution flowed through tygon tubing and an L-shaped glass delivery tube, which were stuffed with nylon mesh to help mixing. The OOE ASW was delivered at 6.3 ml min^{-1} parallel to the axon, and removed from the opposite end and underneath the axon.

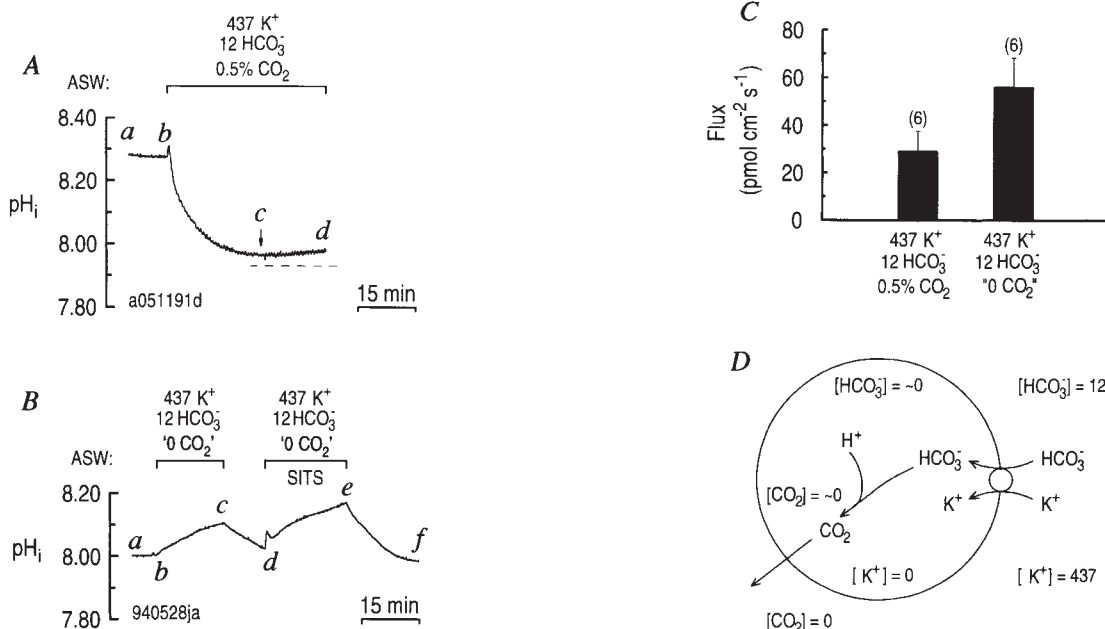


FIG. 2 Comparison of the effects of equilibrated $\text{CO}_2/\text{HCO}_3^-$ ASWs and OOE ASWs containing HCO_3^- (but minimal CO_2) in K^+ -depleted axons. A, Effect of exposing an axon to a pH 8 ASW containing 437 mM K^+ , 12 mM HCO_3^- and 0.5% CO_2 (segment *bcd*). The axon had previously been dialysed to a pH_i of 8.30; dialysis was halted just before point *a*. B, Effect of exposing an axon to an OOE pH 8 ASW containing 437 mM K^+ and 12 mM HCO_3^- but minimal CO_2 (*bc*). The axon had been pre-dialysed to a pH_i of 8.00, to match the pH_i at point *c* in A. During the indicated period, the axon was exposed to 0.5 mM SITS. C, Computed HCO_3^- influx for axons exposed to ASWs containing equilibrated $\text{CO}_2/\text{HCO}_3^-$ (left bar) or OOE HCO_3^- (right bar). The data are taken from experiments like those shown in A and B. D, Model showing how, for a cell exposed to a solution containing HCO_3^- but minimal CO_2 , HCO_3^- cotransported into the cell with K^+ rapidly exits as CO_2 with little intracellular accumulation of CO_2 or HCO_3^- . METHODS. The standard HCO_3^- -free ASW contained no Na^+ , K^+ or Cl^- , and had the following composition (in mM): 437.2 NMDG $^+$, 3.0 Ca^{2+} , 62.5 Mg^{2+} , 563.0 D-glucuronate, 0.1 EDTA $^{2-}$, 5 of the anionic form of

N-(2-hydroxyethyl)piperazine-*N'*-(3-propanesulphonic acid) (EPPS), 5 of the neutral form of EPPS (assuming a pK for EPPS of 8.0) and 0.01 ouabain. In the equilibrated $\text{CO}_2/\text{HCO}_3^-$ ASW, 12 mM D-glucuronate was replaced with 12 mM HCO_3^- . The OOE HCO_3^- ASW was made by mixing (1) a pH 10 solution containing (in mM): 548.9 K^+ , 506.6 D-glucuronate, 16.0 HCO_3^- , 8.0 CO_3^{2-} (assuming a pK for CO_3^{2-} of 10.3), 0.2 EDTA $^{2-}$, 9.9 of the anionic form of EPPS, 0.1 of the neutral form of EPPS, 0.01 ouabain and 0.1 acetazolamide (ACZ) with (2) a pH 6.95 solution gassed with 100% O_2 and containing (in mM): 318.4 K^+ , 6.0 Ca^{2+} , 120.0 Mg^{2+} , 482.0 D-glucuronate, 293.0 mannitol, 0.8 of the anionic form of EPPS, 9.2 of the neutral form of EPPS, 87.6 of the anionic form of 2-(*N*-morpholino)ethanesulphonic acid (MES; assuming a pK of 6.1), 15.1 of the neutral form of MES, 0.01 ouabain and 0.1 ACZ. The dialysis fluid contained no Na^+ , K^+ or Cl^- , and contained (in mM): 417.2 NMDG $^+$, 7.0 Mg^{2+} , 16.0 Tris (hydroxymethyl) aminomethane, 414.0 glutamate, 4.0 ATP $^{4-}$, 1.0 EGTA $^{2-}$, 20.0 HEPES, 95.0 glycine and 0.5 phenol red. The HCO_3^- flux was computed from the rate of pH_i change, β_{Total} (see Fig. 4) and axon diameter 23 .

presence of a new $\text{K}^+/\text{HCO}_3^-$ cotransporter. The out-of-equilibrium approach could be useful in a variety of applications for independently controlling CO_2 and HCO_3^- concentrations and pH.

We generated an out-of-equilibrium (OOE) extracellular solution (artificial sea water, ASW) containing HCO_3^- but very little CO_2 by using a syringe pump to mix rapidly (1) a HCO_3^- solution at pH 10 with (2) a $\text{CO}_2/\text{HCO}_3^-$ -free solution at pH 6.95 (Fig. 1). As the solution passes through the tubing downstream from the mixing 'T', CO_2 is formed: $\text{HCO}_3^- + \text{H}^+ \xrightarrow{\text{fast}} \text{H}_2\text{CO}_3 \xrightarrow{\text{slow}} \text{CO}_2 + \text{H}_2\text{O}$. However, because of the slow reaction, the solution exiting the delivery tube would have only $\sim 3.3 \mu\text{M}$ CO_2 (Fig. 1, legend), compared to $112 \mu\text{M}$ in an equilibrated pH 8 ASW containing 12 mM HCO_3^- . We generated OOE ASWs containing CO_2 but negligible HCO_3^- by mixing a pH 4.95 CO_2 solution with a pH 8.5 $\text{CO}_2/\text{HCO}_3^-$ -free solution. Because of the slow reaction in the sequence $\text{CO}_2 + \text{H}_2\text{O} \xrightarrow{\text{slow}} \text{H}_2\text{CO}_3 \xrightarrow{\text{fast}} \text{HCO}_3^- + \text{H}^+$, the ASW exiting the delivery tube would have only $\sim 122 \mu\text{M}$ HCO_3^- (Fig. 1, legend), compared to 12 mM in an equilibrated pH 8 ASW containing 0.5% CO_2 . All ASWs had a pH of 8.00; OOE solutions contained 0.1 mM acetazolamide to block carbonic anhydrase. We used internal dialysis¹ to control axoplasm composition, and glass microelectrodes² to monitor intracellular pH (pH_i).

Figure 2A shows an experiment in which we dialysed an axon for 80 min with a K^+ - and HCO_3^- -free fluid at pH 8.3, halted dialysis (point *a*) and allowed pH_i to stabilize (*ab*). Introducing an ASW with 437 mM K^+ and 0.5% $\text{CO}_2/12 \text{ mM}$ HCO_3^- caused a rapid pH_i decrease (*bc*). After CO_2 equilibration across the membrane, pH_i increased very slowly (*cd*), presumably due to coupled $\text{K}^+/\text{HCO}_3^-$ uptake^{22,23}. The segment *cd* alkalization requires both K^+ and HCO_3^- (ref. 23). The protocol in Fig. 2B was the same as in 2A except that (1) we dialysed the axon to a pH_i of 8.0, and (2) the ASW containing K^+ and HCO_3^- was a low- CO_2 OOE solution. The pH_i changes caused by this OOE ASW are remarkable in three respects.

First, applying the $\text{K}^+/\text{HCO}_3^-$ ASW produced no initial pH_i decrease (point *b*), further evidence that the OOE ASW contained negligible CO_2 . In some cases, we observed a small pH_i decrease, which averaged 0.027 ± 0.010 ($n = 6$). This value corresponds to an outside CO_2 concentration ($[\text{CO}_2]_o$) of $9 \mu\text{M}$, $\sim 8\%$ of that in an equilibrated ASW.

Second, the pH_i increase evoked by the OOE $\text{K}^+/\text{HCO}_3^-$ ASW (*bc*, Fig. 2B) was much faster than observed with equilibrated ASWs (*cd*, Fig. 2A). The calculated base influx averaged $56 \text{ pmol cm}^{-2} \text{ s}^{-1}$ (Fig. 2C), $\sim$ twofold above the mean of $29 \text{ pmol cm}^{-2} \text{ s}^{-1}$ observed at the same pH_i with equilibrated ASWs²³. This doubling occurred presumably because internal HCO_3^- concentration ($[\text{HCO}_3^-]_i$) was extremely low for axons exposed to the OOE ASW. HCO_3^- transported into the cell would be rapidly converted to CO_2 (Fig. 2D), which would exit into the low- CO_2 ASW. Thus, using an OOE ASW deficient in CO_2 , one can deliver HCO_3^- to the outside of the cell, and yet maintain an unusually low $[\text{CO}_2]_i$ and $[\text{HCO}_3^-]_i$ inside.

Third, removing the OOE ASW caused a monotonic pH_i decrease (*cd*), without the transient pH_i increase expected had the OOE ASW contained appreciable CO_2 (ref. 3). In experiments in which the OOE ASW evoked a small pH_i decrease (see above), the removal led to a transient pH_i increase averaging 0.022 ± 0.006 . This pH_i increase corresponds to a $[\text{CO}_2]_i$ of $\sim 7 \mu\text{M}$, $\sim 6\%$ of that in an equilibrated solution. As known for this $\text{K}^+/\text{HCO}_3^-$ cotransporter²³, the HCO_3^- -induced alkalization was not affected by SITS, 4-acetamido-4'-isothiocyanato-stilbene-2,2'-disulphonate (*de*), which blocks several other HCO_3^- transporters^{4,6}.

The results of Fig. 2 confirm that the $\text{K}^+/\text{HCO}_3^-$ cotransporter mediates HCO_3^- influx into axons dialysed with a K^+ -free fluid. If the cotransporter can mediate HCO_3^- efflux in axons dialysed to a high $[\text{K}^+]_i$ with no K^+ in the ASW, introducing HCO_3^- into the axoplasm should stimulate $\text{K}^+/\text{HCO}_3^-$ efflux and produce a pH_i decrease. To promote a large $\text{K}^+/\text{HCO}_3^-$ efflux, and

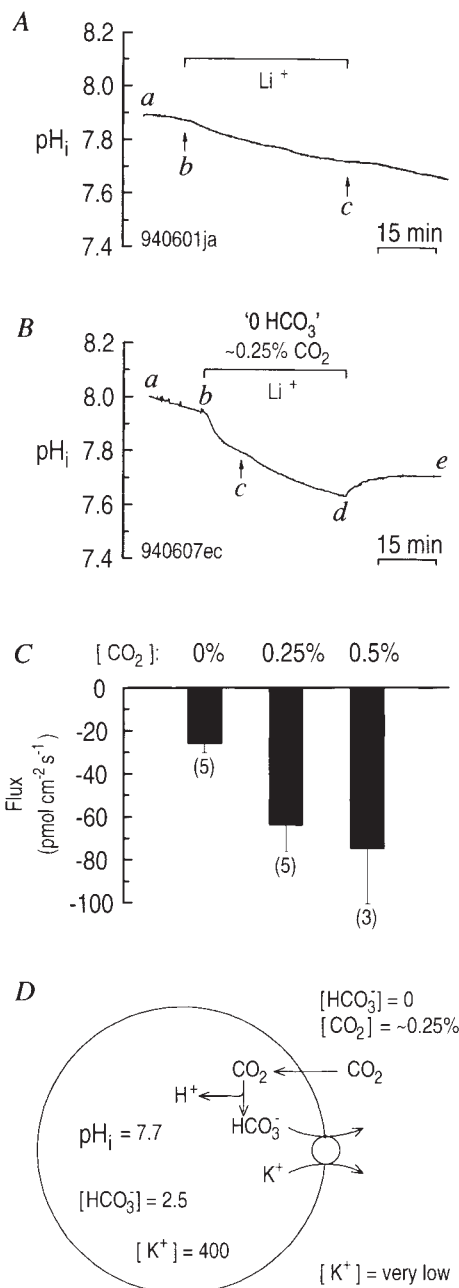


FIG. 3 Comparison of the effects of OOE ASWs containing CO_2 (but minimal HCO_3^-) and $\text{CO}_2/\text{HCO}_3^-$ -free ASWs in K^+ -loaded axons. A, Effect of exposing an axon to a pH 8 ASW containing no CO_2 or HCO_3^- . The axon had been predialysed to a pH_i of 7.79. During the indicated time, the axon was exposed to a Na^+ , K^+ - and Cl^- -free ASW in which the major cation was Li^+ (*bc*). B, Effect of exposing an axon to an OOE ASW containing 0.25% CO_2 but minimal HCO_3^- (*bcd*). The axon had been predialysed to a pH_i of 8.00. The major cation in the OOE solution was Li^+ . C, Computed HCO_3^- efflux for axons exposed to ASWs containing no $\text{CO}_2/\text{HCO}_3^-$ (left bar) or OOE 0.25% CO_2 (middle bar) or OOE 0.5% CO_2 (right bar). The data are taken from experiments like those shown in A and B. D, Model showing how, for a cell exposed to a solution containing CO_2 but minimal HCO_3^- , CO_2 enters the cell and is converted to HCO_3^- , which is then cotransported out of the cell with K^+ .

METHODS. The OOE 0.5% CO_2 ASW was made by mixing (1) a pH 4.95 solution gassed with 1% CO_2 and containing (in mM): 379 Li^+ , 6.0 Ca^{2+} , 120.0 Mg^{2+} , $630.5 \text{ D-glucuronate}$, 0.5 of the anionic form of MES, 9.5 of the neutral form of MES, 0.01 ouabain and 0.1 ACZ with (2) a pH 8.5 solution gassed with 100% O_2 and containing (in mM): 494.0 Li^+ , 410.0 gluconate , 0.2 EDTA^{2-} , 83.6 of the anionic form of EPPS, 26.4 of the neutral form of EPPS, 0.01 ouabain and 0.1 ACZ. The dialysis fluid was similar to that described in Fig. 2, except that 417.2 mM NMDG^+ was replaced by K^+ .

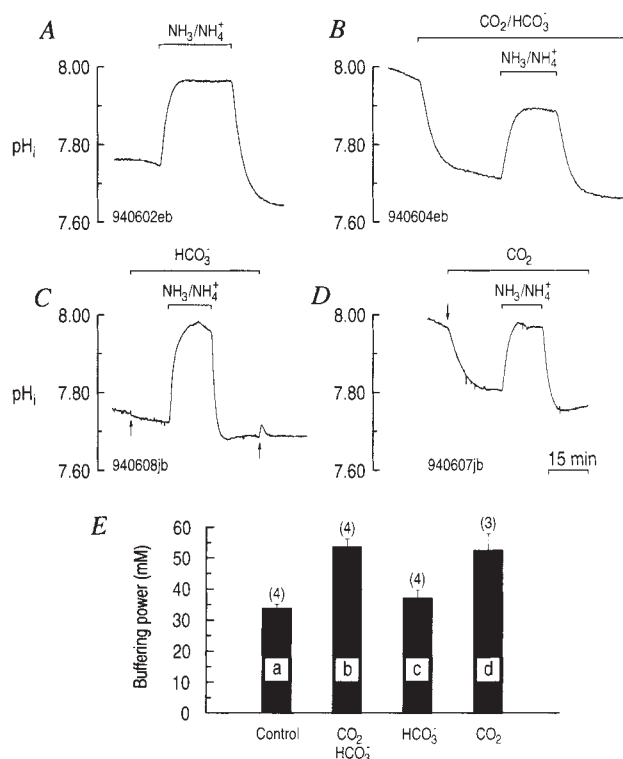


FIG. 4 Effect on intracellular buffering power of exposing axons to equilibrated and OOE CO_2/HCO_3^- solutions. All axons had been pre-equilibrated with a pH 7.75 or a pH 8.00 fluid with an ionic composition similar to that of the fluids used in Fig. 2, except that NMDG⁺ was replaced with tetraethylammonium. The major cation in the ASWs was Li⁺. During the indicated periods, the axons were exposed to a pH 8.00 ASW in which 10 mM Li⁺ was replaced with 10 mM NH_3/NH_4^+ . We computed β from the pH_i decrease elicited by removing the NH_3/NH_4^+ . A, CO_2/HCO_3^- -free ASW. B, Equilibrated 0.5% $CO_2/12$ mM HCO_3^- ASW. Switching to a CO_2/HCO_3^- ASW caused a rapid pH_i decrease, due to CO_2 influx. In the continued presence of CO_2/HCO_3^- , we applied 10 mM NH_3/NH_4^+ . C, OOE 12 mM HCO_3^- (minimal CO_2) ASW. Switching to the OOE HCO_3^- ASW (first arrow) caused only a slight pH_i decrease, indicating the presence of minimal CO_2 in the ASW. In the continued presence of the OOE HCO_3^- , we applied 10 mM NH_3/NH_4^+ . Removing the OOE HCO_3^- (second arrow) caused only a minimal pH_i increase; the small pH_i transient was an electrical artefact. D, OOE 0.5% CO_2 (minimal HCO_3^-) ASW. Switching to the OOE CO_2 ASW (arrow) caused a large pH_i decrease, due to CO_2 influx. In the continued presence of the OOE CO_2 , we applied 10 mM NH_3/NH_4^+ . E, Summary of mean data from experiments like those in A–D. Differences in means are statistically significant for bars a versus b ($P < 0.001$) and for c versus d ($P < 0.05$), but not for bars a versus c or for bars b versus d. The difference between the mean β values in bars b versus a and d versus a is almost exactly the theoretically predicted contribution of an open-system CO_2 buffer to β , that is, $(\ln 10) \times [HCO_3^-]$.

minimize complications of K⁺ in the extracellular unstirred layer⁷, we created a large outwardly directed HCO_3^- gradient by exposing a K⁺-loaded axon to an OOE ASW containing CO_2 but minimal HCO_3^- . We used Li⁺ as the major cation in this OOE solution because it is not a buffer (Fig. 3, legend). The Li⁺ by itself had little effect on pH_i (bc, Fig. 3A). However, an OOE ASW containing Li⁺ and 0.25% CO_2 , but minimal HCO_3^- , caused a rapid initial pH_i decrease (bc, Fig. 3B), due to CO_2 influx, followed by a continuing pH_i decrease (cd), due to K⁺/ HCO_3^- efflux. The mean base efflux in the absence of CO_2 was less than half the efflux in the presence of either 0.25% CO_2 or 0.5% CO_2 (Fig. 3C). Thus, intracellular HCO_3^- stimulates intracellular K⁺-dependent base efflux when $[HCO_3^-]_0$ is negligible (Fig. 3D).

The fluxes in Figs 2C and 3C are computed from the rate of pH_i change and total intracellular buffering power (β_{total}). It has been assumed that β_{total} in the presence of equilibrated CO_2/HCO_3^- is the sum of $\beta_{intrinsic}$ (the β measured in the absence of CO_2) and the theoretical β_{CO_2} (ref. 8). Because some have suggested that CO_2 does not contribute to β (refs 9–12), we directly determined β_{total} from the pH_i decrease caused by removing extracellular NH_3/NH_4^+ (ref. 8) under four conditions: (1) in the absence of CO_2/HCO_3^- (Fig. 4A), where predicted $\beta_{total} = \beta_{intrinsic}$; (2) in the presence of equilibrated 0.5% $CO_2/12$ mM HCO_3^- (Fig. 4B), where predicted $\beta_{total} = \beta_{intrinsic}$

+ β_{CO_2} ; (3) in the presence of an OOE ASW containing 12 mM HCO_3^- but minimal CO_2 (Fig. 4C), where predicted $\beta_{total} = \beta_{intrinsic}$; and (4) in the presence of an OOE ASW containing 0.5% CO_2 but minimal HCO_3^- (Fig. 4D), where predicted $\beta_{total} = \beta_{intrinsic} + \beta_{CO_2}$. Introducing the OOE ASW containing HCO_3^- had no effect on β_{total} , whereas either equilibrated CO_2/HCO_3^- or the OOE ASW containing CO_2 increased β_{total} by almost exactly the theoretical β_{CO_2} (Fig. 4E).

Others^{13–15} have used continuous-flow, rapid mixing to assess H_2CO_3 dehydration, combining HCl and HCO_3^- solutions to produce CO_2/HCO_3^- mixtures with pH values < 5 . We delivered OOE solutions with a physiological pH, and nearly lacking CO_2 or HCO_3^- , to living cells. Although in this study the solutions had the pH and ionic strength of sea water, we have also made comparable OOE solutions for mammalian cells (for example, pH 7.4, 37 °C). OOE solutions could be powerful tools for studying problems other than HCO_3^- transport. For example, respiratory acidosis (a pH decrease caused by elevated blood $[CO_2]$) stimulates respiratory-control centres in the medulla to increase ventilation¹⁶, and stimulates the renal proximal tubule to increase acid secretion^{17,18}. The OOE approach should make it possible to determine unambiguously whether these vital physiological responses are triggered by changes in pH and/or CO_2 , and to determine the sensitivity of chemical reactions to CO_2 , HCO_3^- and pH. □

Received 19 December 1994; accepted 15 February 1995.

1. Brinley, F. J. Jr & Mullins, L. J. *J. gen. Physiol.* **50**, 2303–2311 (1967).
2. Hinkle, J. A. M. in *Glass Electrodes for Hydrogen and Other Cations: Principle and Practice* (ed. Eisenman, G.) 464–477 (Dekker, New York, 1967).
3. Boron, W. F. & De Weer, P. *J. gen. Physiol.* **67**, 91–112 (1976).
4. Cabantchik, Z. I., Knauf, P. A. & Rothstein, A. *Biochim. biophys. Acta* **515**, 239–302 (1978).
5. Thomas, R. C. *Nature* **262**, 54–55 (1976).
6. Boron, W. F. & Boulpaep, E. L. *J. gen. Physiol.* **81**, 53–94 (1983).
7. Frankenhauser, B. & Hodgkin, A. L. *J. Physiol.* **131**, 341–376 (1956).
8. Roos, A. & Boron, W. F. *Physiol. Rev.* **61**, 296–434 (1981).
9. Aickin, C. C. *J. Physiol.* **479**, 301–316 (1994).
10. Baro, I., Eisner, D. A., Raimback, S. J. & Wray, S. J. *J. Physiol.* **417**, 161P (1989).

11. Pocock, G. & Richards, C. D. *Eur. J. Neurosci.* **4**, 136–143 (1992).
12. Roughton, F. J. W. *J. Am. chem. Soc.* **63**, 29–30 (1941).
13. Rossi-Bernardi, L. & Berger, R. L. *J. biol. Chem.* **243**, 1297–1302 (1968).
14. Dalziel, K. *Biochem. J.* **55**, 79–90 (1953).
15. Nattie, E. E. in *Regulation of Breathing, Lung Biology in Health and Disease* Vol. 79 (eds Dempsey, J. A. & Pack, A. I.) 473–510 (Dekker, New York, 1995).
16. Dorman, P. J., Sullivan, W. J. & Pitts, R. F. *J. clin. Invest.* **33**, 82–90 (1954).
17. Brazeau, P. & Gilman, A. *Am. J. Physiol.* **175**, 33–38 (1953).
18. Ho, C. & Sturtevant, J. M. *J. biol. Chem.* **238**, 3499–3501 (1963).
19. Gibbons, B. H. & Edsall, J. T. *J. biol. Chem.* **238**, 3502–3507 (1963).
20. Eigen, M., Kustin, K. & Maass, G. Z. *Physik. Chem.* **30**, 130–136 (1961).
21. Boron, W. F. & Russell, J. M. *J. gen. Physiol.* **81**, 373–399 (1983).
22. Hogan, E. M., Cohen, M. A. & Boran, W. J. *J. gen. Physiol.* (in the press).
23. Hogan, E. M., Cohen, M. A. & Boran, W. J. *J. gen. Physiol.* (in the press).

Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development

Qun-Yong Zhou, Carol J. Quaife
& Richard D. Palmiter

Howard Hughes Medical Institute and Department of Biochemistry,
SL-15, University of Washington, Seattle, Washington 98195, USA

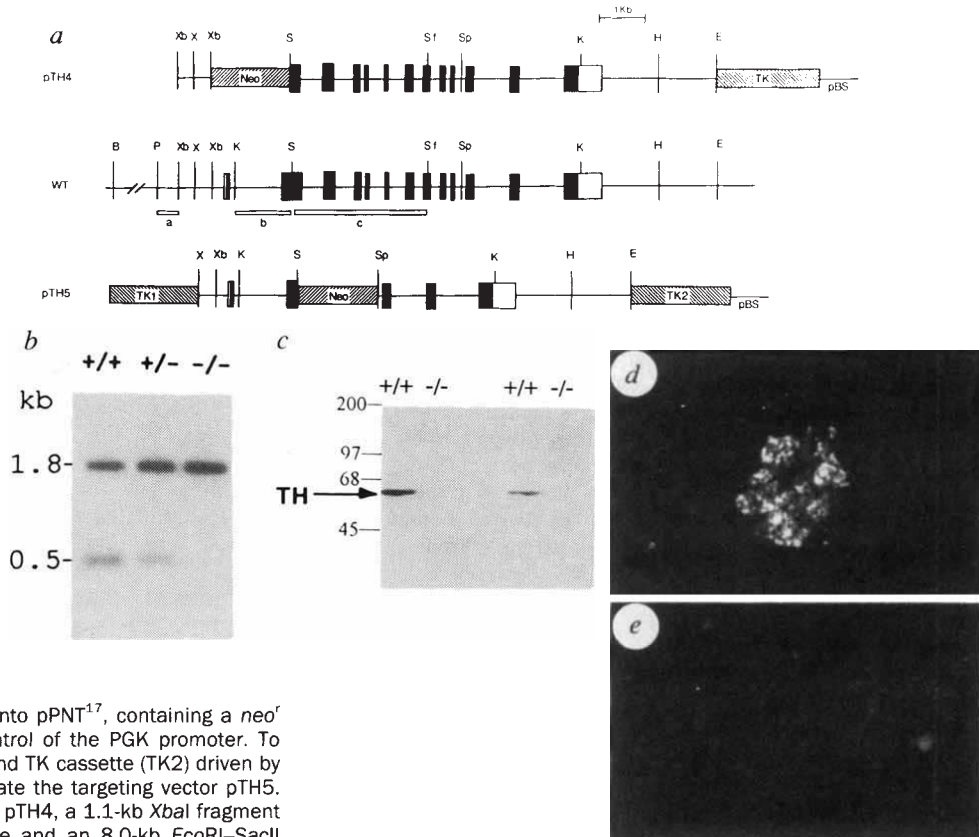
TYROSINE hydroxylase catalyses the initial, rate-limiting step in the catecholamine biosynthetic pathway. Catecholamines, which include dopamine, noradrenaline, and adrenaline, are important neurotransmitters and hormones that regulate visceral functions, motor coordination and arousal in adults¹. The gene encoding tyrosine hydroxylase becomes transcriptionally active in developing neuroblasts during mid-gestation of rodent embryos, before the onset of neurotransmission²⁻⁶. Here we show that inactivation of both tyrosine hydroxylase alleles results in mid-gestational lethality: about 90% of mutant embryos die between embryonic days

11.5 and 15.5, apparently of cardiovascular failure. Administration of L-DOPA (dihydroxyphenylalanine), the product of the tyrosine hydroxylase reaction, to pregnant females results in complete rescue of mutant mice *in utero*. Without further treatment, however, they die before weaning. We conclude that catecholamines are essential for mouse fetal development and postnatal survival.

The organization of the tyrosine hydroxylase (TH) gene and two targeting vectors that were used to generate null mutations of the TH gene are shown in Fig. 1a. The pTH5 construct replaces two thirds of the coding region with a neomycin-resistance gene (*neo^r*), whereas the pTH4 construct replaces the proximal promoter and first two exons with the *neo^r*. Both targeting vectors include thymidine kinase genes to allow selection against random integration⁷. Linearized targeting constructs were electroporated into AB1 embryonic stem (ES) cells and the cells were selected with G418 and ganciclovir. We identified 3 out of 144 and 11 out of 315 ES cell clones with predicted genomic structures for targeting vectors pTH5 and pTH4, respectively. Selected ES cell clones were microinjected into C57BL/6J blastocysts and transferred into the uteri of pseudopregnant recipient females. Two independent ES cell clones derived from targeting vector pTH5 and three independent ES cell clones derived from targeting vector pTH4 were transmitted through the germ line. heterozygous mice from all ES cell clones grew normally and showed no obvious anatomical or behavioural defects. Heterozygotes were crossed to generate TH homozygous (*TH^{-/-}*) mice. Reverse-transcribed polymerase chain reaction (RT-PCR) ana-

FIG. 1 a, Targeting of the mouse TH gene in embryonic stem cells and mice. Genomic structure and restriction map of the mouse TH gene locus and targeting vectors pTH4 and pTH5 used to disrupt the TH gene are shown. Black boxes represent exons. Probes used for Southern and dot blot analysis are shown as white bars. B, *Bam*HI; H, *Hind*III; K, *Kpn*I; P, *Pst*I; S, *Sac*II; Sf, *Sfi*I; Sp, *Spe*I; X, *Xho*I; Xb, *Xba*I. b, RT-PCR analysis of TH mRNA from E12.5 embryos. c, Western blot analysis of TH protein from neonatal brains. d, e, Glyoxylic acid-induced catecholamine (CA) histochemistry of wild-type (d) and mutant adrenal gland from P15 mice (e).

METHODS. The mouse TH gene was cloned from a 129 Sv/CPJ genomic library by standard colony hybridization procedure using a rat TH cDNA as a probe. An 11-kb genomic DNA fragment containing 2.5 kb of 5' flanking sequence, the entire coding region and 1.5 kb of 3' flanking sequence was subcloned into pBS (Stratagene). A 2.5-kb *Xho*I–*Sac*II fragment and a 4-kb *Spe*I–*Eco*RI fragment were inserted into pPNT¹⁷, containing a *neo^r* and a TK cassette (TK1) under the control of the PGK promoter. To increase the targeting frequency, a second TK cassette (TK2) driven by its own promoter was inserted to generate the targeting vector pTH5. To construct the second targeting vector pTH4, a 1.1-kb *Xba*I fragment of 5' flanking sequence of the TH gene and an 8.0-kb *Eco*RI–*Sac*II fragment of TH sequence were inserted into pPNT. The *neo^r* cassettes in both constructs are oriented in the same direction as the TH gene. The unique *Not*I sites in both constructs allowed the plasmids to be linearized. AB1 ES cells were cultured, electroporated, and selected as described¹⁸. For pTH5, colonies were picked and plated on new feeder cells and expanded for freezing and Southern blot hybridization. For pTH4, a third of each colony was plated on new feeder cells and the remainder was subjected to PCR analysis with oligonucleotides for the PGK promoter (5'-CCAGACTGCCTTGGGAAAG-3') and the TH gene (5'-GTAAGTACAGCAGG-CAGTTG-3'). PCR-positive colonies were confirmed



by Southern blot hybridization using probe a. A *neo^r* probe was also used to ensure that there was only one copy of each construct in the ES cell clones. Chimaeras were produced by microinjecting about 15 ES cells into C57BL/6J blastocysts and transplanting the embryos into the uteri of C57BL/6J × CBA F1 pseudopregnant females. Chimaeric males were tested for germ-line transmission of the agouti coat phenotype of 129Sv-derived ES cells. For RT-PCR, total RNA was extracted

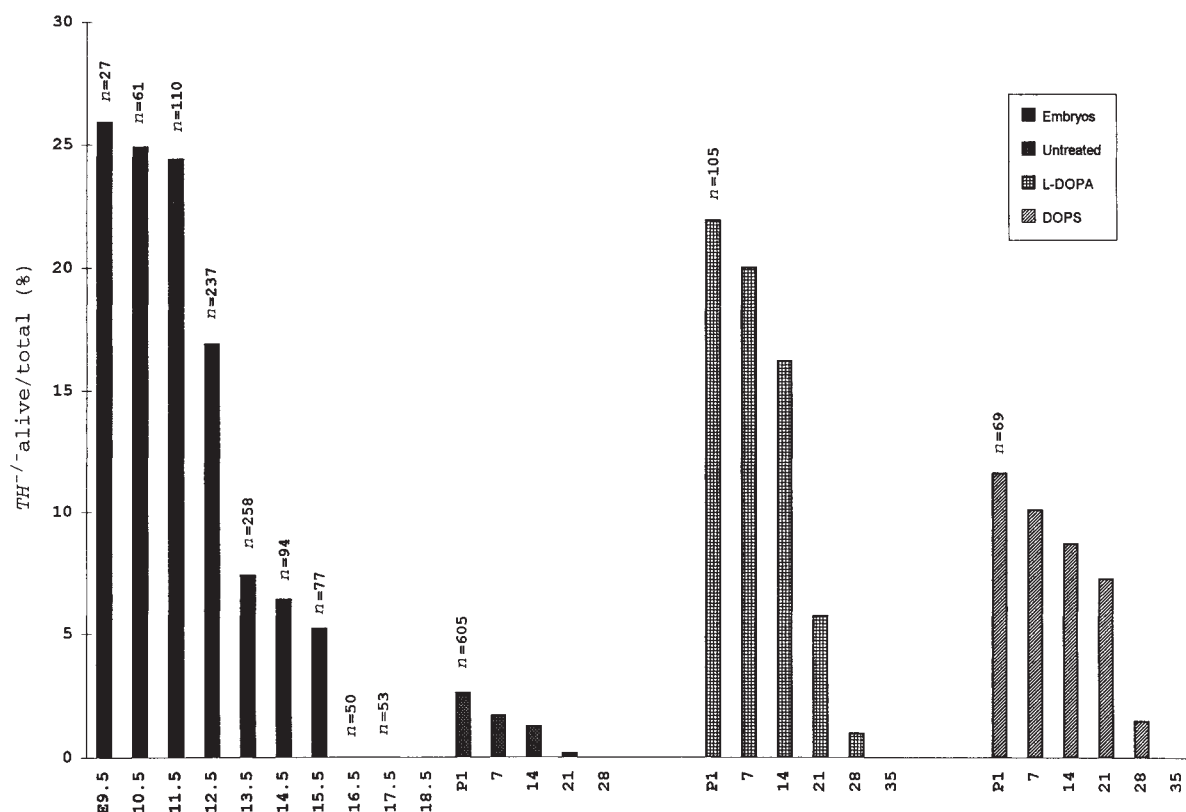


FIG. 2 Percentage of live $TH^{-/-}$ fetuses and offspring from untreated, L-DOPA and DOPS-treated heterozygous females. The numbers of fetuses and offspring genotyped are shown at the top of the bars. The number shown on the bars labelled P1 indicate pups born during the preceding 24 h. For drug treatment, heterozygotes were mated. The day on which a vaginal mucous plug was observed was considered E0.5. Females were given L-DOPA (1.0 mg ml^{-1}) or DOPS (0.5 mg ml^{-1}) in drinking water beginning at E8.5 of pregnancy. Drugs were prepared daily in water containing 0.25% ascorbic acid and the water bottles

were protected from light. To genotype mice or embryos, DNA was prepared from the tail or the yolk sac. Southern blot hybridizations were done using probe a or c (Fig. 1a). For routine genotyping, mutants were identified by the presence of the *neo*^r sequence and absence of the TH gene by dot hybridization using a *neo* probe and probe b or c for targeting vector pTH4 and pTH5, respectively (Fig. 1a). Alternatively, the absence of TH gene was determined by PCR with primers that lie within the deletions.

lysis showed the absence of TH transcripts in $TH^{-/-}$ mutants (Fig. 1b). The absence of TH protein in $TH^{-/-}$ mutants was demonstrated by western blotting using a monoclonal anti-TH antibody (Fig. 1c). Glyoxylic acid-induced catecholamine histo-fluorescence (Fig. 1e) was greatly reduced in mutants compared

to controls (Fig. 1d). These observations suggest that these mutations completely inactivate TH enzymatic activity.

No $TH^{-/-}$ mutants were found out of first 23 pups screened, implying that TH-deficient mice die *in utero*. To investigate the timing of the lethality, mice from embryonic days E9.5 to E17.5 were genotyped. Of the viable embryos, 25% were $TH^{-/-}$ at E11.5, but by E15.5 that number was reduced to 5% and only 2.6% of the mice born were $TH^{-/-}$ (Fig. 2). The same embryonic lethality was observed for embryos derived from both targeting vectors.

To ascertain why the $TH^{-/-}$ embryos die *in utero*, 14 mutants (E12.5 ($n=4$), E13.5 ($n=6$), E14.5 ($n=1$) and E15.5 ($n=3$)) and 9 controls were prepared for histological examination. Examination of haematoxylin and eosin-stained paraffin sections revealed that the major organs were normal with the exceptions noted below. Congestion of blood in the liver and major vessels was noticed in 6 of 14 mutants (compare Fig. 3a, b; c, d). In 2 of 3 transversely sectioned E12.5 mutant embryos, the atria were markedly dilated. In many areas, the walls of atria were reduced to one or two cell layers (Fig. 3f). This degree of thinning of the atrial wall was not seen in any of the controls (Fig. 3e). Closer examination of the heart showed subtle morphological differences between mutants and mice with at least one intact TH gene. In many regions of mutant hearts, ventricular cardiomyocytes were more heterogeneous in size and appeared less organized (Fig. 3h) than in controls (Fig. 3g). In addition, mutant embryos showed slight bradycardia. When counted in a saline buffer at room temperature, the heart rates of exercised E12.5 embryos were 37.9 ± 7.5 beats per minute ($n=7$, s.d.) for

from E12.5 wild-type and mutant embryos. Oligonucleotides OTH 7 (5'-TTTGAC-CCCAGACACAGCA-3'), and OTH8 (5'-GGTAGAATACAG CATGAA-3') were designed from sequences of exon 12 and exon 13¹⁹, which are intact in both targeting vectors. RT-PCRs were done as described²⁰. The 0.5 kb band is the predicted RT-PCR product. The 1.8-kb band represents contaminating genomic DNA in the RNA preparations. A faint 0.5-kb band was also visible in mutant lanes after longer exposure, which may be due to the low-frequency read through of transcripts initiated within the *neo*^r cassette, because the *neo*^r cassette is in the same orientation as the TH gene. For western blotting, neonatal brains were lysed in a buffer containing 150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 1 mM phenylmethylsulphonyl fluoride. An equal volume of $2 \times$ SDS sampling buffer was added and the samples were clarified by boiling and centrifugation. Protein (30 μ g) was loaded onto a 10% SDS-polyacrylamide gel. Following electrophoresis, protein was transferred to a nitrocellulose filter, the filter was rinsed with TST buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween 20) and blocked for 1 h at room temperature in TSTM (TST+5% non-fat dried milk). Tyrosine hydroxylase was identified by incubating the blot first with mouse monoclonal antibody against tyrosine hydroxylase (IncStar), then washing in TST and incubating with horseradish peroxidase-conjugated anti-mouse antibody in TST. After extensive washing, the filter was developed using Enhanced Chemiluminescence Detection (Amersham). Glyoxylic acid CA histo-fluorescence was performed as described²¹.

mutants versus 44.9 ± 7.0 ($n=25$, s.d.) for either heterozygous or wild-type embryos ($t=2.080$ (Student's t -test), $P<0.05$). The blood congestion, bradycardia, and unusual histology of the cardiomyocytes suggest that catecholamines are necessary for proper cardiovascular function during midgestation, but which cell type(s) respond directly to catecholamines is unknown.

To test whether the absence of catecholamines is responsible for the embryonic fatality of $TH^{-/-}$ mutants, we attempted to rescue $TH^{-/-}$ *in utero* with L-DOPA, the product of TH enzymatic activity. L-DOPA was supplied in the drinking water of pregnant females, starting at E8.5 of pregnancy. L-DOPA at 0.25 mg ml^{-1} rescued about 50% of the mutants to E17.5 or

birth (data not shown), whereas 1 mg ml^{-1} rescued all $TH^{-/-}$ mutants to birth (Fig. 2). These data provide compelling evidence that the lack of catecholamines is responsible for the embryonic fatality of $TH^{-/-}$ mutants. The continuous presence of catecholamines seems to be necessary because treatment of pregnant females with 1 mg ml^{-1} L-DOPA from E10.5 to E13.5 resulted in only partial rescue of $TH^{-/-}$ *in utero* (data not shown). The rescue by L-DOPA rules out the possibility that metabolism of an unrecognized substrate by TH is essential for embryonic development or that a chromosomal perturbation introduced at the TH locus disrupts the normal regulation of nearby developmentally important genes such as *H19* or *IGF2* (refs 8–10).

Targeted disruption of the dopamine β -hydroxylase (DBH) gene revealed that noradrenaline is essential for fetal survival¹¹. To assess whether dopamine is also essential for embryonic development, we attempted to rescue TH mutants *in utero* with D,L-threo-3,4-dihydroxyphenylserine (DOPS), a substrate that can be converted to noradrenaline but not dopamine by the action of L-aromatic-amino-acid decarboxylase¹². At 0.5 mg ml^{-1} , a dose that completely rescued *DBH*^{-/-} mice *in utero*¹¹, DOPS only partially rescued $TH^{-/-}$ mice to birth (Fig. 2), indicating that, in addition to noradrenaline, dopamine may also be essential for fetal development.

What is the source of embryonic catecholamines? TH-positive neuroblasts are first apparent by immunocytochemistry in the developing brain and in the cells migrating from the neural crest around E10 in fetal mouse^{2,6,13–14}, and TH messenger RNA can be detected as early as E8.5 by RT-PCR¹¹. In the brain, TH-positive neuroblasts will give rise to the dopaminergic neurons of midbrain and the noradrenergic neurons of the brainstem. In the periphery, the neural-crest-derived cells give rise to sympathetic neurons, adrenal chromaffin cells, enteric neurons, and small intensely fluorescent cells of sympathetic ganglia and paraganglia. All of these TH-positive neuroblasts are potential sources of catecholamines. Because most of the TH-deficient embryos die before the time that catecholaminergic neurotransmission is established¹⁵, catecholamines probably act in a paracrine manner but it is not known whether their secretion is regulated. Inactivation of the *MASH1* gene resulted in mouse embryos with few TH-positive sympathetic neurons but they survived to term¹⁶. Presumably, residual sympathetic neurons and other TH-positive neuroblasts provide sufficient catecholamines.

Of 1,184 offspring that were born of C57BL6/129 hybrids, 30 were $TH^{-/-}$, implying incomplete penetrance of embryonic lethality in the outbred background (Fig. 2); however, no $TH^{-/-}$ mutants have been identified out of 205 inbred 129Sv/CPJ offspring. We suspect that maternal catecholamines account for the survival of a new $TH^{-/-}$ mice to term¹¹. Mutant mice that survived to term were born of normal weight but runting became apparent in all mutants by postnatal day (P)7. At P15, mutants weighed $39 \pm 8\%$ ($n=9$, s.d.) of non-mutant littermates and they appeared very weak. Unlike controls, they could not balance on a rotating pencil or walk up an incline. All TH-deficient mice succumbed by 4 weeks of age (Fig. 2).

Like the few naturally born $TH^{-/-}$ mice, all rescued $TH^{-/-}$ pups were severely runted, displayed defective motor function, and died within 5 weeks (Fig. 2). In contrast, about 40% of *DBH*^{-/-} pups survived to adulthood¹¹. The death of all $TH^{-/-}$ pups before reaching adulthood indicates that dopamine is

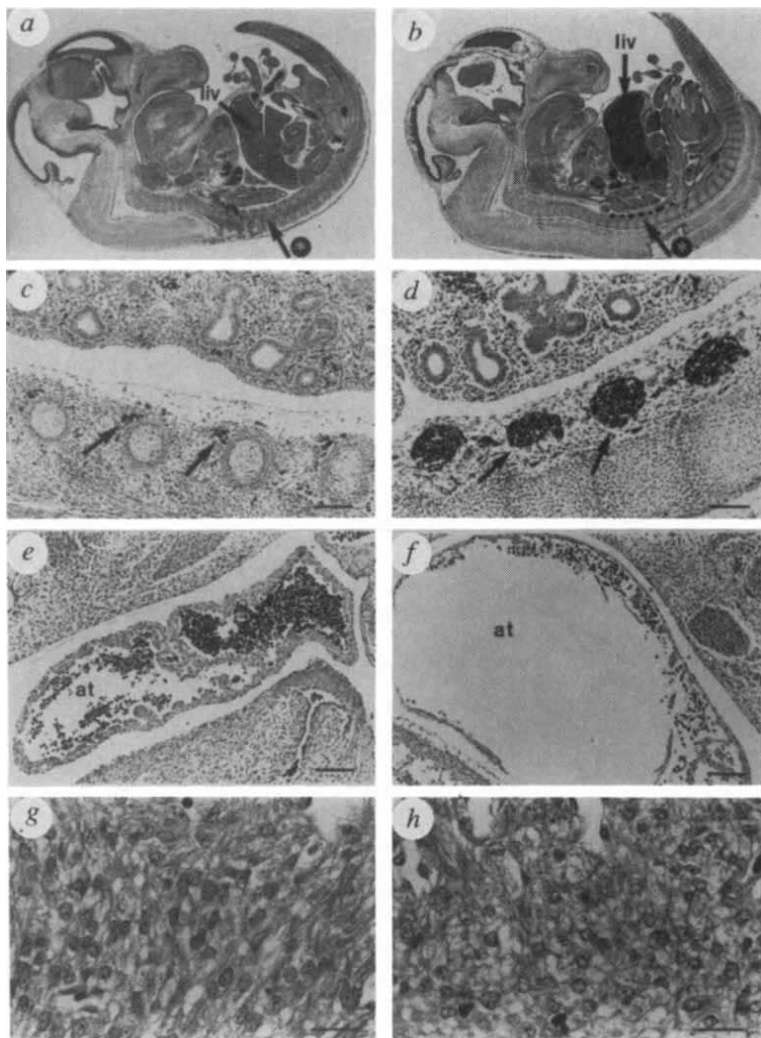


FIG. 3 Histological anomalies associated with the TH gene disruption. A comparison of sagittal sections from a day E13.5 wild-type fetus (a) and $TH^{-/-}$ mutant (b). The liver (liv) of the mutant embryo (b) is congested with blood. The arrows marked with asterisks in a and b indicate regions shown at higher magnification in c and d. Vertebral blood vessels (arrows) in the wild-type (c) are compared with the congested vessels in the mutant fetus (d). Atria (at) of a wild-type E12.5 (e) and mutant fetus (f). Note the unusual expansion of the atrial chamber in this mutant. Cellular organization of the ventricle of a E12.5 wild-type fetus (g) compared with a $TH^{-/-}$ littermate (h). One of the regions of a mutant ventricle in which the cells appear more vacuolated and less organized is shown in h. Scale bar, $100 \mu\text{m}$ (c–f); $20 \mu\text{m}$ (g, h).

METHODS. Pregnant females were killed by CO_2 inhalation. Fetuses were isolated and their viability was determined by the presence of a heart beat. After removal of yolk sac, fetuses were fixed in 10% neutral formalin for at least 24 h, dehydrated with gradient ethanol and embedded in paraffin. Sections ($5 \mu\text{m}$) were obtained and stained with haematoxylin and eosin. Mouse developmental stage was determined according to ref. 22.

essential for postnatal survival. The defective motor function of TH-deficient pups probably reflects deficiency of the nigrostriatal motor control pathway¹. It will be important to examine the consequence of TH deficiency on the development of catecholaminergic systems of the central nervous system. □

Received 26 October 1994; accepted 3 February 1995.

- Cooper, J. R., Bloom, F. E. & Roth, R. H. *The Biochemical Basis of Neuropharmacology* 221–337 (Oxford Univ. Press, New York, 1991).
- Cochard, P., Goldstein, M. & Black, I. B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1986–1990 (1978).
- Teitelman, G., Baker, H., Joh, T. H. & Reis, D. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 509–513 (1979).
- Jonakait, G. M., Rosenthal, M. & Morrell, J. I. *J. histochem. Cytochem.* **37**, 1–5 (1989).
- Teitelman, G., Gorshon, M. D., Rothman, R. P., Joh, T. H. & Reis, D. J. *Dev. Biol.* **86**, 348–355 (1981).
- Specht, L. A., Pickel, V. M., Joh, T. H. & Reis, D. J. *J. comp. Neurol.* **199**, 233–253 (1981).
- Mansour, S. I., Thomas, K. R. & Capecchi, M. R. *Nature* **336**, 348–352 (1988).
- O'Malley, K. L. & Rotstein, P. *Nucleic Acids Res.* **16**, 4437–4446 (1988).
- Bartolomei, M. S., Zemel, S. & Tilghman, S. M. *Nature* **351**, 153–155 (1991).
- DeChiara, T. M., Efstratiadis, A. & Robertson, E. J. *Nature* **345**, 78–80 (1990).
- Thomas, S. A., Mutsumoto, A. M. & Palmiter, R. D. *Nature* **374**, 643–646 (1995).
- Kato, T., Karai, N., Katsuyama, M., Nakamura, M. & Katsube, J. *Biochem. Pharmac.* **36**, 3051–3057 (1987).
- Hall, A. K. & Landis, S. C. *J. Neurosci.* **11**, 472–484 (1991).
- Gershon, M. D., Rothman, T. P., Joh, T. H. & Teitelman, G. N. *J. Neurosci.* **4**, 2269–2280 (1984).
- Rubin, E. J. *Neurosci.* **5**, 673–684 (1985).
- Guillemot, F. et al. *Cell* **75**, 463–476 (1993).
- Tybulwicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).
- Soriano, P., Montgomery, C., Geske, R. & Bradley, A. *Cell* **64**, 693–702 (1991).
- Iwata, K. et al. *Biochem. biophys. Res. Commun.* **182**, 348–352 (1992).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbour Laboratory Press, New York, 1989).
- De La Torre, J. C. *J. Neurosci. Res.* **3**, 1–5 (1980).
- Kaufman, M. H. *The Atlas of Mouse Development* (Academic, San Diego, 1992).

ACKNOWLEDGEMENTS. We greatly appreciate the technical assistance from L. Chow, K. Clegg, T. Dawe, G. Froelich and S. Thomas. We thank A. Berns for mouse genomic library; A. Bradley and P. Soriano for ES cells; E. Lewis for rat TH cDNA probe; R. Mulligan for pPNT vector; O. Civelli for providing carbidopa and our colleagues for comments on the manuscript. This study was supported in part by a grant from the NIH.

Noradrenaline is essential for mouse fetal development

Steven A. Thomas*, Alvin M. Matsumoto† & Richard D. Palmiter*

*Department of Biochemistry, SL-15, Howard Hughes Medical Institute, and †Gerontology Research Education and Clinical Center, ZB-20, Veterans Administration Medical Center, and Department of Medicine, University of Washington School of Medicine, Seattle, Washington 98195, USA

CATECHOLAMINES such as noradrenaline and adrenaline have been implicated in numerous physiological processes^{1–4} but, although catecholamine synthesis begins at mid-gestation⁵, previous studies have provided little evidence for any role in early development^{6,7}. Furthermore, there are several case reports of humans with noradrenaline deficiency⁸. To investigate this, we used gene targeting⁹ to produce mice lacking dopamine β-hydroxylase and therefore unable to synthesize noradrenaline or adrenaline. We report here that in heterozygous mothers, most homozygous embryos died *in utero*, and only about 5% reached adulthood. Survival probably depends on catecholamine transfer across the placenta because, in homozygous mothers, all embryos die *in utero*. Mortality was due to lack of noradrenaline *in utero* because it could be prevented by treatment with dihydroxyphenylserine, a precursor that can be converted to noradrenaline in the absence of dopamine β-hydroxylase. Mutant embryos had a histological phenotype similar to that of embryos deficient in tyrosine hydroxylase¹⁰, suggesting that death might be due to cardiovascular failure.

To disrupt the dopamine β-hydroxylase (DBH) locus, 3.4 kilobases (kb) of genomic sequence that included 0.6 kb of the

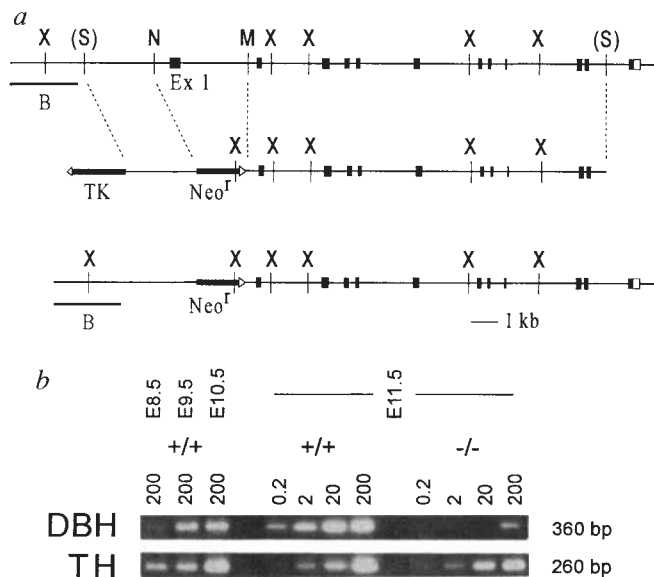
DBH proximal promoter and the first exon were replaced with a neomycin-resistance cassette (*neo'*, Fig. 1a), and the resulting construct was electroporated into AB1 embryonic stem (ES) cells. Of 84 clones selected, 7 contained a targeted disruption of the DBH gene, and six of these clones were each injected into C57BL/6J blastocysts. Five clones gave rise to chimaeras, and four transmitted the disrupted allele through the germ line. Heterozygotes appeared normal in all respects, and were inbred to generate homozygous (*DBH*^{−/−} or mutant) offspring. The genotypes of the first 75 offspring were 18 wild-type mice, 56 heterozygotes, and one homozygote, indicating that most homozygotes died *in utero*. To determine when homozygous fetuses die, females from heterozygous crosses were killed at specific times during gestation. All 20 mutant fetuses examined at E10.5 were alive, whereas only 56 and 26% of mutant fetuses at E11.5 and E18.5, respectively, were alive (Fig. 2a). In contrast, 97% of non-mutant littermates (*n* = 74) were alive at E13.5. Because homozygous fetuses accounted for 25% of total fetuses (*n* = 102) at E11.5, it is unlikely that homozygotes were dying before E11. Genotyping of all offspring (*n* = 1,375) born from heterozygous crosses indicates that 12% of *DBH*^{−/−} mice survive to term. One of the four lines was maintained on the inbred 129/SvCPJ background, and a similar frequency of mutant mice was born (Fig. 2a). Of those that survived to term, 40% died within 48 h of birth, and another 20% died between the third and fifth postnatal weeks (Fig. 2b). Mutant mice could not be distinguished from littermates shortly after birth; however, runting was usually apparent by postnatal day 7 (P7). By P21, mutant mice weighed ~50% of their littermate controls. *DBH*^{−/−} mice eventually grew to 80% (males, s.e.m. = 2, *n* = 24) and 88% (females, s.e.m. = 2, *n* = 24) of littermate adult weight, but the growth phase that normally occurs between 3 and 5 weeks was delayed by one to two weeks in the mutants (Fig. 3). Because of the runting, mutant mice were weaned around P30 rather than P20. Adult mutant mice were morphologically normal, showed no increased mortality, exhibited normal mobility, and reacted in the same manner as littermates to handling. Besides the runting, the only other distinguishing feature of the mutant mice was ptosis.

On the basis of the design of the targeting construct (Fig. 1a), there should be no messenger RNA for DBH in the knockout mice. To test this, and determine when during development DBH and tyrosine hydroxylase (TH) mRNAs are first expressed, sensitive reverse-transcribed polymerase chain reaction (RT-PCR) assays were developed. A moderate signal for both TH and DBH mRNA was detected from whole-fetus RNA at the earliest time examined, E8.5, and a considerably stronger signal for both was detected at E9.5 and later. Whole-fetus RNA from E11.5 litters of heterozygote crosses was used to examine knockouts. TH mRNA was detected from all fetuses using 0.2 ng of RNA (Fig. 1b). Likewise, DBH mRNA was detected from all wild-type and heterozygous fetuses using 0.2 ng of RNA with primer pairs located at either end of the mRNA. However, no signal could be detected for the 5'-end of DBH mRNA in any of the knockout fetuses using 0.2 to 200 ng of RNA, which was expected because this region had been deleted by gene targeting. Similarly, no signal could be detected for the 3'-end of DBH mRNA in any of the knockout fetuses using 0.2 to 20 ng of RNA, but a signal was detected with 200 ng. This may be due to a low frequency of transcripts initiated within *neo'* that extend into the 3'-end of the DBH gene, which has the same transcriptional orientation as *neo'*.

To assess whether noradrenaline and adrenaline are synthesized in the mutant mice, and to determine the levels of dopamine, catecholamines were measured using high-performance liquid chromatography (HPLC) with electrochemical detection. Dopamine and noradrenaline were first detected at E9.5 and E10.5, respectively, whereas adrenaline was consistently detected in all fetuses older than E13.5 (Fig. 4a). At E11.5, *DBH*^{+/−} fetuses had a small but significant reduction (19%, *P* < 0.02) in noradrenaline and an increase (103%, *P* < 0.2) in dopamine (Fig.

FIG. 1 Genetic construct used to disrupt the murine *DBH* allele, and detection of mRNA for *DBH* and *TH* using RT-PCR. **a**, Boxes indicate the location of the 12 exons, the first of which is labelled Ex 1. The top diagram represents the native *DBH* allele; the middle diagram the targeting construct (*DTK*); and the bottom diagram the disrupted *DBH* allele. For *DTK*, the proximal promoter and the first exon were replaced with *neo^r* and the herpes simplex virus thymidine kinase gene (*TK*) was added to the 5'-end. **b**, The developmental expression of *DBH* (3'-end) and *TH* mRNA from E8.5 to E11.5 using 200 ng of total RNA. In addition, detection of mRNA is compared between mutant and control fetuses at E11.5 using 0.2 to 200 ng of total RNA.

METHOD. A λ -clone was isolated from a C57BL/6J mouse genomic library using a PCR-amplified fragment of exon 1 as a probe. The λ -clone was subcloned into the *Sall* (S) site of pBluescript (Stratagene). The proximal promoter and the first exon (3.4 kb) were excised using unique *NarI* (N) and *MluI* (M) sites, and *neo^r* was inserted between these sites. *TK* was added to the 5'-end of the *DBH* construct (*DTK*). *DTK* (20 μ g) was linearized with *NotI* and electroporated in the presence of 10^7 AB1 ES cells (230 V, 500 pF pulse followed immediately by a 240 V, 500 pF pulse). Positive/negative selection was achieved using 350 μ g ml^{-1} of G418 and 2 μ M ganciclovir, respectively. DNA from ES cell colonies was digested with *XbaI* (X) and screened for targeting by Southern blot hybridization with a 2.5-kb genomic *DBH* *BstXI*–*BstEII* fragment (B) just 5' of the targeting construct *DBH* sequence. The native *DBH* allele gave two bands of 8.5 and 3.0 kb, whereas the targeted *DBH* allele gave two bands of 5.4 and 3.0 kb. Chimaeric mice were generated by injecting 15–30 targeted ES cells into C57BL/6J blastocysts 3.5 days postcoitus. Injected blastocysts were transferred into pseudopregnant C57/B6 $\times$ CBA F_1 hybrids 2.5 days postcoitus. Chimaeras with >80% agouti coat colour were bred with C57BL/6J mice to test for germ-line transmission. Some chimaeras positive for transmission also were bred with 129/SvCPJ mice. For detection of mRNA, fetuses were dissected and frozen on dry ice. Instruments were washed between each fetus at E11.5 to prevent cross-contamination. Tissue was processed with a Brinkman homogenizer in 5 M guanidine isothiocyanate and 8% v/v β -mercaptoethanol buffered with 50 mM EDTA and 10 mM Tris, pH 7.5. The homogenizer was washed in 3 half-litre bottles of autoclaved dH_2O between samples. RNA was precipitated with 4 M LiCl, redissolved, extracted with phenol:chloroform (1:1), and quantified spectrophotometrically. Reverse transcription (RT) was carried out using 200 ng of RNA per μ l of buffer, which consisted of 75 mM KCl, 4.5 mM MgCl_2 , 0.25 mM CaCl_2 , 100 μ g ml^{-1} bovine serum albumin (Sigma), 5 mM dithiothreitol, 0.3 units μl^{-1} RNAGuard (Pharmacia), 0.1 units μl^{-1} reverse transcriptase (Life Sciences), 200 μ M dNTPs (Pharmacia), 1 μ M oligonucleotide, 20 mM Tris, pH 8.3. Primer oligonucleotides for RT were 5'-GTCAACCTTGGTGACCAAGTGGTGA (exon 4, mTH), 5'-AGATGGATCTGCCCTTTCCGG (exon 2, mDBH) and 5'-ATCTGCCTCTGTGTGATGGGC (exon 12, mDBH). Primers for PCR were: for mTH, 5'-ATGGAAATGCTGTTCTCAACCTGC (exon 2, sense) and 5'-ACCCCTGTCTCTCTGGCACTGC (exon 3, antisense); for mDBH 5'-end (not shown), 5'-GCACTGCTGTGGCCATCTTCTGGTC (exon 1, sense) and 5'-GCCCTTTCCGGTCACTCCAGGC (exon 2, antisense); and for mDBH 3'-end, 5'-GGAGGAGATGTGTGCAACTATGTGC (exon 10, sense) and 5'-GGGCTCTGAGTCTGTGGATGG (exon 12, antisense). PCR was done in 25 μ l of buffer containing 1 μ l of either the RT mixture or 10-fold dilutions of the RT mixture. The PCR buffer was: 16.6 mM ammonium sulphate, 6.7 mM MgCl_2 , 67 mM Tris, pH 8.8, 1 mM dNTPs, 100 μ g ml^{-1} BSA, 5 units ml^{-1} Taq polymerase (Perkin Elmer) and 1 μ M oligonucleotides.



4b), suggesting that *DBH* becomes rate-limiting in E11.5 *DBH^{+/-}* fetuses. In *DBH^{+/-}* fetuses from heterozygous females, greatly reduced levels of noradrenaline (3% of those for wild-type fetuses) and dramatically elevated levels of dopamine (17-fold) were detected at E11.5. One potential source of the remaining noradrenaline was placental transfer of maternal catecholamines^{11,12}. Unlike offspring of heterozygous females, noradrenaline was not detected in the *DBH^{+/-}* fetuses of homozygous females. In addition, fetal lethality was more severe in homozygous mothers: 67% of the *DBH^{+/-}* fetuses had died by E11.5 (Fig. 2a). However, 18% of *DBH^{+/-}* fetuses from homozygous females were also dead at E11.5, indicating that homozygous females are less reproductively fit than heterozygous females. This difference could account for the greater mortality of *DBH^{+/-}* fetuses in homozygous females. Furthermore, none of the pups born ($n=66$) to homozygous females mated with heterozygous males have been *DBH^{+/-}*. These results indicate that there is no detectable *DBH* activity in the mutant mice, and suggest that catecholamines may cross the placenta to rescue some fetuses in heterozygous females. Because noradrenaline levels are normally much higher than those of adrenaline from E10.5 to E13.5 (Fig. 4a), noradrenaline rather than adrenaline probably activates adrenergic receptors during early fetal development.

During the dissection of the fetuses, we were unable to discern any differences between living mutant mice and their littermates. Fetuses from E10.5 (2), E12.5 (2), E13.5 (8) and E15.5 (1) were processed for histology. Serial sagittal sections were stained with haematoxylin and eosin for examination. In one E13.5 mutant

fetus, pooling of blood in the liver and large vessels was observed. In addition, the myocardiocytes appeared disorganized and atrophied, whereas other organs appeared histologically normal. In the other fetuses examined, no gross morphological abnormalities were observed between mutant and non-mutant littermates. However, greater heterogeneity in cell size and cellular orientation in mutant hearts, similar to the changes observed in fetuses lacking *TH*¹⁰, was observed in 8/11 fetuses E12.5 to E15.5. Mutant fetuses from homozygous females mated to heterozygous males were also examined histologically at E10.5. Three of the six mutant fetuses also had increased heterogeneity of myocardiocytes, which was not observed in the two heterozygous littermates.

To test whether the absence of noradrenaline is responsible for the fetal lethality in *DBH*-deficient mice, pregnant heterozygous females were treated with dihydroxyphenylserine (DOPS), which can be converted into noradrenaline by L-aromatic-amino-acid decarboxylase (AADC)¹³. Fresh drinking water containing DOPS was provided daily starting at E8.5. Figure 2c shows that there was a dose-dependent rescue of the mutant mice. Essentially all homozygotes survived to term at a dose of 1 mg ml^{-1} . When carbidopa, an inhibitor of AADC, was included with the DOPS, rescue was no longer seen, indicating that this effect of DOPS was dependent on its conversion to noradrenaline. Carbidopa did not reduce the number of non-mutant neonates born, indicating that it did not have a nonspecific toxic effect. Catecholamine levels were measured in E11.5 fetuses from heterozygous females treated with 1 mg ml^{-1} DOPS in their drinking water beginning at E8.5. Noradrenaline levels were 10%

(s.e.m. = 6, $n=4$) and adrenaline levels were 82% (s.e.m. = 8, $n=4$) of those for heterozygous fetuses from untreated females measured simultaneously.

These data show that noradrenaline is essential for early fetal development and suggest that the heart may be an important target for its action. The absence of striking morphological changes in most cases suggests that the primary defect may be physiological, for example an inability to maintain sufficient heart rate or contractility. Hypoxia stimulates

release of fetal catecholamines¹⁴, and catecholamines protect neonates during hypoxia¹⁵. Perhaps noradrenaline provides a feedback mechanism so that the heart can maintain adequate perfusion during the rapid fetal growth that occurs between E8.5 and E11.5.

In general, our results are very similar to those for $TH^{-/-}$ mice¹⁰; however, some differences are apparent. Significant levels of fetal mortality occur in $DBH^{-/-}$ mice one day earlier than in $TH^{-/-}$ mice. Also, 59% of $DBH^{-/-}$ mice from heterozygous

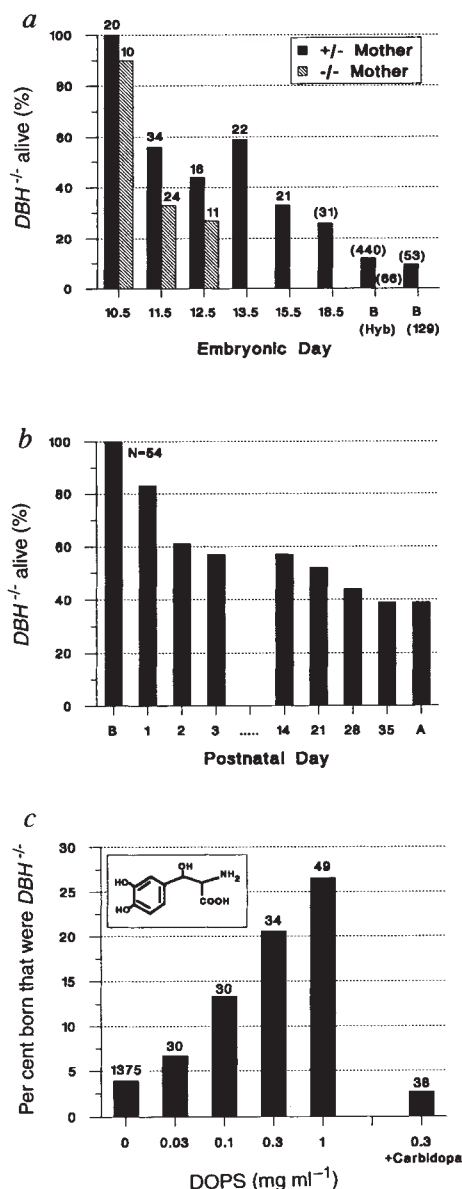


FIG. 3 Postnatal growth of DBH mutant mice and their littermates. Squares represent the average weight of 6 littermates ($\pm$ s.e.m.), and other symbols represent individual mutant mice. Two of the four mutant mice died during the fourth postnatal week (indicated by a cross). All mice were female except for one of the mutants that died at 3 weeks.

METHODS. Pups were weighed once a week beginning on the day of birth (except one mutant, which was begun on the fifth week). Coat colour and sex were used to distinguish pups until they were ear-tagged and genotyped (third postnatal week).

FIG. 2 Survival of $DBH^{-/-}$ mice and the effect of DOPS therapy. **a**, Prenatal survival of $DBH^{-/-}$ fetuses from heterozygous (solid bars) and homozygous (hatched bars) females. Numbers over bars are the total number of mutant fetuses examined, and numbers in parentheses are estimates of the number expected (see methods). Hyb indicates data from 129/SvCPJ $\times$ C57BL/6 hybrids, and 129 indicates data from inbred 129/SvCPJ mice. **b**, Postnatal survival of $DBH^{-/-}$ mice from heterozygous females, with the total number that were born (B) being 54. Adult mice (A) are over 8 weeks old. **c**, Dose effect for rescue of mutant fetuses by supplying DOPS to pregnant females from E8.5 to E18.5. Because offspring are from heterozygous parents, complete rescue should result in ~25% $DBH^{-/-}$ neonates. The dose of 1 mg ml^{-1} DOPS resulted in 30% mortality of non-homozygous neonates, whereas there was no mortality of non-homozygous neonates with 0.3 mg ml^{-1} . Where indicated, carbidopa (0.1 mg ml^{-1}) was included with the DOPS. The biochemical structure of DOPS is shown; decarboxylation of DOPS by AADC produces noradrenaline.

METHODS. Male and female heterozygotes were housed together, and every morning the females were checked for mucous vaginal plugs. The day a plug was found was considered to be E0.5. The timing of pregnancy was verified for excised fetuses by their crown-rump lengths according to ref. 18. Genotyping was done using double dot-hybridization analysis of nucleic acid from tissue digested with proteinase K (Boehringer-Mannheim). Dots were probed with the *neo*^r gene and a fragment of the *DBH* gene that had been deleted from the genome. At E18.5 we were unable reliably to genotype resorbing fetuses that had died around E11. Thus the total number of mutant fetuses was estimated to be 25% of the total number of fetuses and resorptions observed at this stage. At birth, the expected number of $-/-$ mice from heterozygous crosses was estimated as one third of the total number of $+/+$ and $+/-$ mice. The expected number of $-/-$ mice from homozygous females crossed with heterozygous males was estimated to be equal to the number of $+/+$ mice born. P1 is the first day that neonates were observed, at which time some had already died, but these were counted as being born (B) alive although some may have been stillborn. Fresh DOPS (D,L-threo-3,4-dihydroxyphenylserine; Research Biochemicals) was prepared daily in water containing 2 mg ml^{-1} ascorbic acid (to retard oxidation), and the solution was protected from light. Carbidopa was from Research Biochemicals.

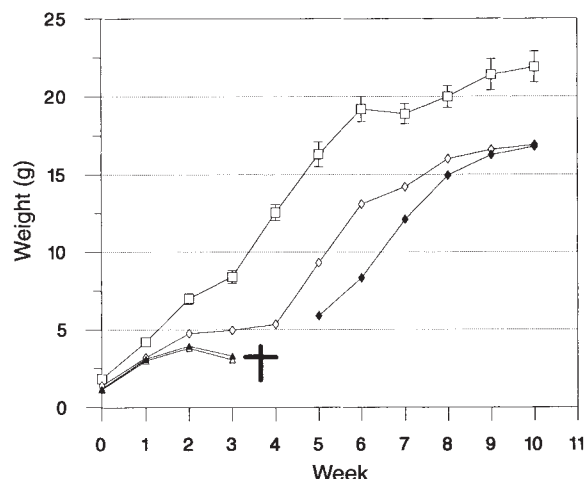
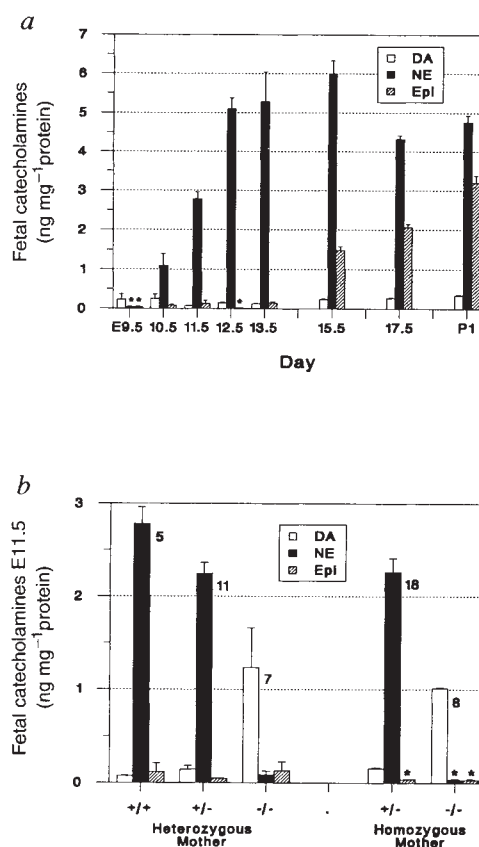


FIG. 4 Quantification of catecholamines in wild-type and mutant fetuses. **a**, Developmental appearance of catecholamines in whole wild-type fetuses. Each time point is the average of 5 fetuses except for E9.5, which was the average of 3 pooled litters. **b**, Catecholamines in mutant fetuses and littermates at E11.5 from heterozygous and homozygous pregnant females. Numbers above the bars indicate number of fetuses. Asterisks (*) indicate that all samples were below the limit of detection.

METHODS. Catecholamines were measured using reverse-phase ion-pairing HPLC with amperometric electrochemical detection using a modification of methods described previously¹⁹. Briefly, fetuses were dissected, frozen on dry ice and stored at -70°C until processed. The fetal yolk sac was used for genotyping. Fetuses E13.5 and older were homogenized (Janke & Kunkel Ultra-Turrax T25) on ice in buffer (8 ml g^{-1} tissue for E13.5, 4.5 ml g^{-1} for E15.5 and older). A 400 μl sample was then sonicated on ice (Branson 450). Fetuses E12.5 and younger were directly sonicated on ice in 200 μl of buffer. The homogenization/sonication buffer was 0.1 M perchloric acid containing 0.01% cysteine and 5 $\mu\text{g l}^{-1}$ of 3,4-dihydroxybenzylamine (internal standard). Protein was determined by Bio-Rad protein microassay using either 10 μl (E9.5–E12.5) or 50 μl (E13.5–P1) of sonicate. Sonicates were centrifuged for 10 min, 18,700g at 4°C . The supernatant (or a 200 μl aliquot for E13.5–P1) was added to 20 mg of acid-washed alumina and 800 μl of 0.5 M Tris with 2% EDTA, pH 8.0. The mixture was rotated overnight at 4°C . Following a brief spin, the pelleted alumina was washed twice with 1 ml of ice-cold double-distilled H_2O . Catecholamines were extracted with 100 μl (E9.5–E12.5) or 200 μl (E13.5–P1) of 0.1 M perchloric acid containing 0.01% cysteine. Samples were stored at 4°C overnight. For HPLC, the mobile phase was composed of 0.1 M citrate-phosphate buffer, 0.1 mM EDTA, 0.012% octanesulphonic acid and 7% methanol, at pH 3.8, and was run at a flow rate of 1 ml min^{-1} . The injection volume was 20 μl for standards and samples, except for samples from E9.5–E12.5, where it was 40 μl . The working potential of the electrochemical detector (BAS LC-4C) was $+0.8\text{ V}$, and the full scale sensitivity was 2 nA (E9.5–E12.5) or 5 nA (E13.5–P1). Quantification of catecholamines was done by comparing the peak heights of unknowns to those of known quantities of catecholamine standards and DHBA using an HP 3393A integrator. The lower limit of detection for dopamine, noradrenaline and adrenaline was 5 pg per 20 μl .



females are still alive at E13.5, whereas only 30% of $TH^{-/-}$ mice are alive then. Both of these differences are statistically significant ($P < 0.05$). One explanation for these differences is that high levels of dopamine, present in $DBH^{-/-}$ but not $TH^{-/-}$ mice, may affect normal adrenergic functions. Activation of adrenergic or dopaminergic receptors by dopamine could either exacerbate or partially alleviate the lack of adrenergic receptor stimulation by noradrenaline. This difference suggests that dopamine may also play a critical role during fetal development. Another possibility is that *p*-octopamine, which can be synthesized from tyrosine by the sequential actions of AADC and DBH, contributes to fetal heart development. *p*-Octopamine

would be produced in $TH^{-/-}$ but not $DBH^{-/-}$ mice. *p*-Octopamine has been found in the normal adult mouse heart¹⁶ and the developing rat heart and hypothalamus¹⁷.

The death of $DBH^{-/-}$ mice *in utero* is somewhat surprising, not only because experimental lesions to the adrenergic nervous system are not life-threatening, but also because at least 6 patients have been documented to be deficient in DBH, most probably because of genetic defects⁸. The results in mice suggest that these people may be the lucky few who survived this condition. $DBH^{-/-}$ mice offer a new approach to understanding the role of noradrenaline and adrenaline in development, physiology and behaviour. □

Received 26 October 1994; accepted 3 February 1995.

1. Cannon, W. B. & Rosenbluth, A. *Autonomic Neuroeffector Systems* (Macmillan, New York, 1937).
2. Levi-Montalcini, R. & Angeletti, P. U. *Pharmac. Rev.* **18**, 619–628 (1966).
3. Berti, F., Lentati, R. & Usardi, M. M. *Med. pharmac. Exp.* **13**, 227–232 (1965).
4. Angeletti, P. U. & Levi-Montalcini, R. *Proc. natn. Acad. Sci. U.S.A.* **65**, 114–121 (1970).
5. Schlumpf, M., Lichtensteiger, W., Shoemaker, W. J. & Bloom, F. E. in *Biogenic Amines in Development* (eds Parvez, H. & Parvez, S.) 567–590 (Elsevier/North Holland Biomedical Press, Amsterdam, 1980).
6. Rosen, T. S., Lin, M., Spector, S. & Rosen, M. R. *J. pharmac. exp. Ther.* **208**, 118–122 (1979).
7. Kudlacz, E. M., Navarro, H. A., Eylers, J. P. & Slotkin, T. A. *J. dev. Physiol.* **13**, 243–249 (1990).
8. Roberts, D. et al. *Hypertension* **18**, 1–8 (1991).
9. Cappelletti, M. R. *Science* **244**, 1288–1292 (1989).

10. Zhou, Q.-Y., Quaife, C. J. & Palmiter, R. D. *Nature* **374**, 640–643 (1995).
11. Morgan, C. D., Sandler, M. & Panigel, M. *Am. J. obstet. Gynecol.* **112**, 1068–1075 (1972).
12. Parvez, S. & Parvez, H. *Hormone Res.* **5**, 321–330 (1974).
13. Kato, T., Karai, N., Katsuyama, M., Nakamura, M. & Katsube, J. *Biochem. Pharmac.* **36**, 3051–3057 (1987).
14. Jones, C. T. & Robinson, R. O. *J. Physiol.* **248**, 15–33 (1975).
15. Seidler, F. J. & Slotkin, T. A. *J. Physiol.* **358**, 1–16 (1985).
16. Baldessarini, R. J. *J. Neurochem.* **18**, 2509–2518 (1971).
17. David, J. C., Cavo, A., Coulon, J. F. & Delacour, J. *Neuroscience* **12**, 1271–1276 (1984).
18. Theiler, K. *The House Mouse: Atlas of Embryonic Development* (Springer, New York, 1989).
19. Liebmann, J. E. & Matsumoto, A. M. *Endocrinology* **126**, 555 (1990).

ACKNOWLEDGEMENTS. We thank G. Cadd for some of the early cloning work; M. Appleby, E. Brandon and R. Idzerda for helpful instruction with blastocyst injections; R. Perlmutter for the use of the microinjection equipment; A. Bradley and P. Soriano for supplying the AB1 ES cells; C. Quaife for help with histological analysis, and T. Dawe, G. Froelick, W. Green and B. Marck for their technical support. This work was supported in part by grants from the NIH.

Activation of microglial cells by β -amyloid protein and interferon- γ

Lucia Meda^{*†}, Marco A. Cassatella^{*}, Gyorgyi I. Szendrei[‡], Laszlo Otvos Jr[‡], Pierluigi Baron[†], Martin Villalba[§], Davide Ferrari[§] & Filippo Rossi^{*||}

^{*} Institute of General Pathology, University of Verona, 37134 Verona, Italy

[†] Institute of Neurology, University of Milano, 20122 Milano, Italy

[‡] Wistar Institute, Philadelphia, Pennsylvania 19104, USA

[§] Institute of General Pathology, University of Ferrara, 44100 Ferrara, Italy

^{||} To whom correspondence should be addressed

ALZHEIMER'S disease is the most common cause of progressive intellectual failure¹. The lesions that develop, called senile plaques, are extracellular deposits principally composed of insoluble aggregates of β -amyloid protein (A β), infiltrated by reactive microglia and astrocytes^{2,3}. Although A β , and a portion of it, the fragment 25–35 (A β (25–35)), have been shown to exert a direct toxic effect on neurons^{4,6}, the role of microglia in such neuronal injury remains unclear⁷. Here we report a synergistic effect between A β and interferon- γ (IFN- γ) in triggering the production of reactive nitrogen intermediates and tumour-necrosis factor- α (TNF- α) from microglia. Furthermore, using co-culture experiments, we show that activation of microglia with IFN- γ and A β leads to neuronal cell injury *in vitro*. These findings suggest that A β and IFN- γ activate microglia to produce reactive nitrogen intermediates and TNF- α , and this may have a role in the pathogenesis of neuronal degeneration observed in ageing and Alzheimer's disease.

To test whether the interaction of A β with microglia could induce the production of proinflammatory and potentially cytotoxic mediators, we measured the accumulation of NO₂⁻ from cultures of N9 microglial cells⁸ stimulated with A β (25–35), the full-length A β (1–42) and, for comparison, with IFN- γ and lipopolysaccharide (LPS). As shown in Fig. 1, IFN- γ alone was a stimulus of NO₂⁻ accumulation⁸, but doses of A β (25–35) ranging from 10 to 150 μ g ml⁻¹, as well as LPS⁸, did not induce detectable NO₂⁻ accumulation, even after a 72-h incubation (not

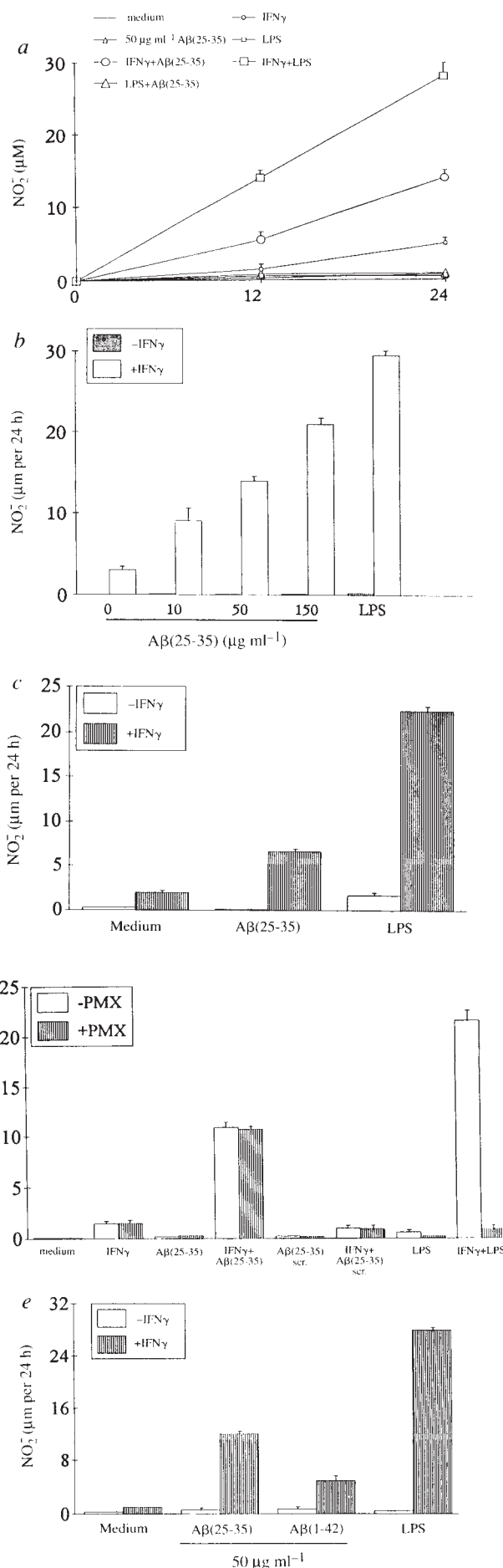


FIG. 1 Effect of A β peptides on the accumulation of NO₂⁻ from murine microglia. Mean values $\pm$ s.d. of assays performed with supernatants collected from triplicate wells for various conditions are shown. **a**, A representative time-course experiment ($n=4$). **b**, A representative dose-response experiment ($n=3$). **c**, A representative experiment performed with primary microglia ($n=3$). **d**, The effect of A β (25–35) scrambled (scr.) peptide and of polymyxin-b sulphate (10 μ g ml⁻¹) ($n=3$). **e**, The effect of A β (1–42) ($n=3$).

METHODS. Murine microglial cell lines N9 (ref. 8) and primary microglial cell cultures (>98% Mac-1 positive) prepared from two-day-old postnatal Swiss CD1 mouse cerebral cortex as previously described²¹, were plated in 96-well tissue-culture plates (Nunc, Roskilde, Denmark) at 4×10^4 per 100 μ l per well and incubated in the presence or absence of various doses of A β (25–35)s, 50 μ g ml⁻¹ A β (1–42), 100 ng ml⁻¹ LPS, and/or 100 U ml⁻¹ IFN- γ . After the indicated times, microglial cell supernatants were assayed for NO₂⁻ accumulation by the Griess reaction, according to methods described previously²². Peptides [A β (1–42), A β (25–35), GSNKGAIIGLM, and A β (25–35) scrambled, IMLKGNASIG, (both with free amino termini and amidated carboxy terminal) were synthesized, purified and characterized as described previously²³ and resuspended in sterile phosphate buffered saline. Recombinant human and murine IFN- γ were obtained from G. Garotta (LaRoche, Basel, Switzerland). Polymyxin-b sulphate (PMX), lipopolysaccharide (LPS, from *Escherichia coli* 026:B6), and all other reagents were from Sigma. Solutions used throughout our experiments were prepared with endotoxin-free water for clinical use²⁴.

shown). However, $A\beta(25-35)$ potentiated the accumulation of NO_2^- from microglia induced by $IFN-\gamma$, in a time- (Fig. 1a) and dose-dependent manner (Fig. 1b). Similar results were obtained by using another $A\beta(25-35)$ peptide (Sigma) and after stimulation of a different microglial clone (N13)⁸, as well as with three different preparations of primary mouse (Fig. 1c) and rat microglia (not shown). Specificity of the effects of $A\beta(25-35)$ was demonstrated by the fact that a peptide with the same amino-acid composition as $A\beta(25-35)$, but with a scrambled sequence, did not influence NO_2^- accumulation from microglia treated with $IFN-\gamma$ or LPS (Fig. 1d). Furthermore, any possible influence of trace levels of contaminating LPS in our $A\beta(25-35)$ stocks was excluded because polymyxin-b sulphate (PMX) almost completely abolished the accumulation of NO_2^- triggered by $IFN-\gamma$ plus LPS, but failed to affect NO_2^- accumulation triggered by

$IFN-\gamma$ alone, or by $IFN-\gamma$ plus $A\beta(25-35)$ (Fig. 1d). Figure 1e shows that, as observed for $A\beta(25-35)$, $A\beta(1-42)$ by itself had no effect on this response, but when $A\beta(1-42)$ was used in combination with $IFN-\gamma$, it induced significant NO_2^- accumulation. The effect of $A\beta(1-42)$ was dose and time dependent, and was not blocked by PMX (not shown). The effects of $A\beta(1-42)$ on NO_2^- accumulation and the responses described below were observed only if the stock solutions of the peptide were preincubated for 7 days at 37 °C before use. Preliminary experiments showed that both $A\beta(1-40)$ and $A\beta(1-42)$ proteins were ineffective in inducing NO_2^- accumulation, and in potentiating the action of $IFN-\gamma$, if they were freshly added to the microglia cultures at doses up to 50 $\mu g\ ml^{-1}$. These results are consistent with previous observations^{6,9} showing that synthetic $A\beta$ peptides can either have very little cytotoxicity or produce profound cell injury, depending on time-related conformational transitions and aggregation state during their handling or storage in a variety of buffers.

Subsequently, we investigated whether $A\beta(25-35)$ and/or $IFN-\gamma$ were also able to induce the production of $TNF-\alpha$ by microglial cells. As shown in Fig. 2, untreated cells constitutively produced very small amounts of $TNF-\alpha$, whereas microglia stimulated with $A\beta(25-35)$, but not with the scrambled $A\beta(25-35)$ (not shown), released significant levels of $TNF-\alpha$, in a dose- (Fig. 2a) and time-dependent manner (maximal production after 12 h; not shown). Co-stimulation of microglia with $A\beta(25-35)$ and $IFN-\gamma$ resulted in an additive effect on $TNF-\alpha$ production (Fig. 2a), which was not influenced by the presence of PMX (Fig. 2b). Similar results were also obtained with primary murine microglia cells (Fig. 2c). To determine whether the endogenous $TNF-\alpha$ produced in response to $A\beta(25-35)$ could contribute to the observed NO_2^- accumulation¹⁰, we performed studies in which neutralizing antibodies to $TNF-\alpha$ were added to microglia stimulated with $A\beta(25-35)$, LPS, or $TNF-\alpha$, in the presence or absence of $IFN-\gamma$ (Fig. 3). Anti- $TNF-\alpha$ antibodies, over a 24-h time course, only slightly ($8 \pm 2\%$, mean $\pm$ s.d., $n=3$) affected the accumulation of NO_2^- stimulated by LPS plus $IFN-\gamma$, whereas that induced in response to $A\beta(25-35)$ plus $IFN-\gamma$ was reduced by $64 \pm 4\%$. As a control, anti- $TNF-\alpha$ antibodies completely suppressed the accumulation of NO_2^- stimulated by $TNF-\alpha$ plus $IFN-\gamma$, whereas an isotype-matched antibody had no effect in any of these conditions. Thus a large part of the NO_2^- accumulation induced by $A\beta(25-35)$ plus $IFN-\gamma$ is mediated by endogenous $TNF-\alpha$.

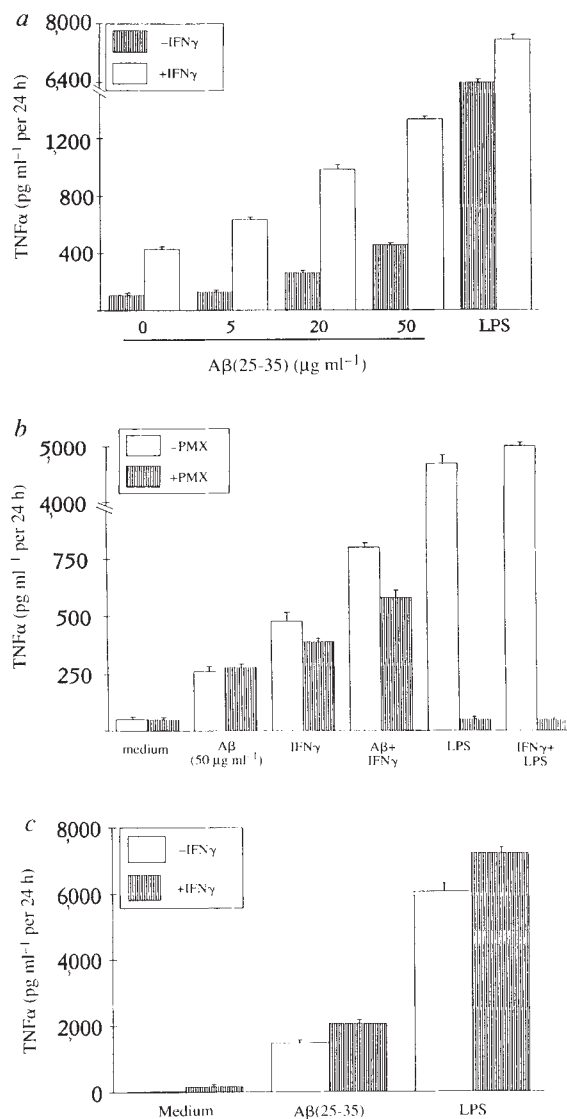


FIG. 2 Effect of $A\beta(25-35)$ on $TNF-\alpha$ release by murine microglia. Mean values $\pm$ s.d. of triplicate assays performed with supernatants collected and pooled from three wells. a, b, Representative experiments with N9 microglia cell lines. c, Representative experiment with primary murine microglia cells.

METHODS. Microglial cells were cultured in 96-well plates and stimulated with different concentrations of $A\beta(25-35)$ and LPS, in the presence or absence of $IFN-\gamma$ (a and c) ($n=3$), or with $A\beta(25-35)$ and LPS, with or without $IFN-\gamma$ and in the presence or absence of PMX (b) ($n=2$). Cell-free supernatants were collected, and antigenic mouse $TNF-\alpha$ was determined using an ELISA kit (Factor-Test-Xtm mouse $TNF-\alpha$; Genzyme), with a detection limit of 30 $pg\ ml^{-1}$.

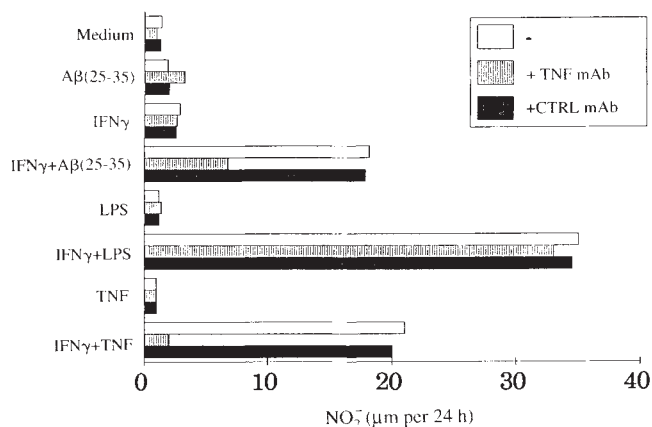
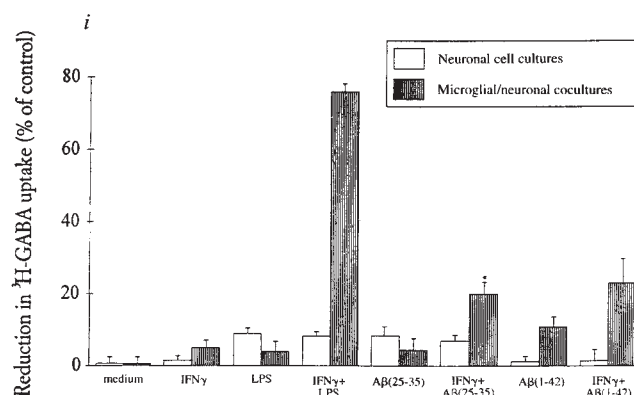
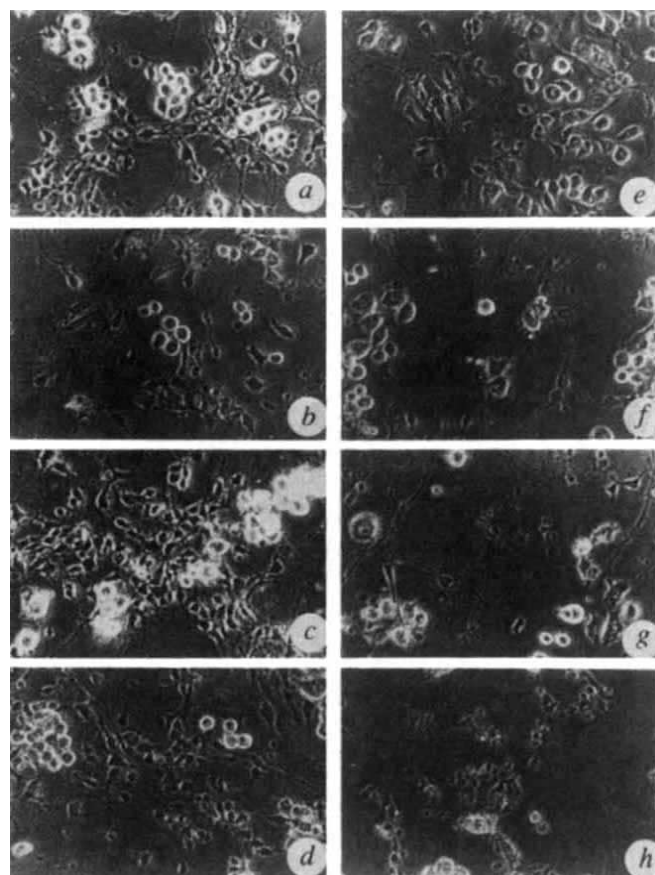


FIG. 3 Effect of anti- $TNF-\alpha$ antibody on $A\beta(25-35)$ plus $IFN-\gamma$ -stimulated NO_2^- accumulation. The experiment shown is representative of three. **METHODS.** N9 microglia were stimulated with 50 $\mu g\ ml^{-1}$ $A\beta(25-35)$, 100 $ng\ ml^{-1}$ LPS, or 100 $U\ ml^{-1}$ $TNF-\alpha$ in the presence or absence of $IFN-\gamma$ (100 $U\ ml^{-1}$), with or without an anti- $TNF-\alpha$ antibody (XT22 (ref. 25), 20 $\mu g\ ml^{-1}$), or control isotype-matched (CTRL) antibody. At the 24-h timepoint, supernatants were collected for NO_2^- measurement.

FIG. 4 Morphological appearance of primary cultures of neurons co-incubated with microglia and exposed to $A\beta(25-35)$, $A\beta(1-42)$, LPS, in the presence or absence of $IFN-\gamma$ (a-h); and 3H -GABA uptake by neuronal cells in microglial-neuronal cell co-cultures stimulated with $A\beta$ s or LPS, in the presence or absence of $IFN-\gamma$ (i). Phase-contrast micrographs of neuronal cultures co-incubated for 72 h with microglia in the absence (a-d) or presence (e-h) of $IFN-\gamma$, and after treatment with diluent (a, e), $50 \mu g ml^{-1}$ $A\beta(25-35)$ (b, f), $50 \mu g ml^{-1}$ $A\beta(1-42)$ (c, g) and $100 ng ml^{-1}$ LPS (d, h). Note that in a-e no morphological evidence of neuronal injury is visible; in f and g marked neuronal degeneration is observable, that is, cell debris, somal shrinkage, neuritic swelling and regression; in h morphological evidence of dramatic neuronal damage is evident, that is, loss of neuronal cells, disappearance of neurites and cell debris.

METHODS. Primary cultures of neuronal cells isolated from rat brain cortex were prepared essentially as described previously²⁶. Neuronal cells were plated at 7.5×10^4 cells cm^{-2} and maintained in culture in Dulbecco's modified Eagle's medium with 2 mM glutamax and 10% horse serum. Co-cultures were then prepared by adding 5×10^4 microglial cells to five-day neuronal cell cultures. They were treated immediately with the indicated doses of $A\beta$ peptides or LPS, in the presence or absence of $IFN-\gamma$. Micrographs are representative of at least four independent experiments, each performed in triplicate. For the 3H -GABA uptake experiments (i) microglial-neuronal cell co-cultures were incubated for 48 h with $50 \mu g ml^{-1}$ $A\beta(25-35)$, $50 \mu g ml^{-1}$ $A\beta(1-42)$, $100 ng ml^{-1}$ LPS, in the presence or absence of $100 U ml^{-1}$ $IFN-\gamma$, before determining 3H -GABA uptake (as described previously²⁷). The bars represent the percentage of inhibition of 3H -GABA uptake in each condition over the corresponding control (mean $\pm$ s.e. of at least four experiments performed in triplicate): 100% corresponds to $80 fmol min^{-1}$ per 10^5 cells. Statistical significance was assessed using Student's *t*-test for paired data. Asterisks indicate $P < 0.025$ versus stimuli alone in both co-cultures and neuronal cell cultures. There was no statistically significant difference between any other conditions and their respective controls.



Finally, to investigate whether microglia activation by $A\beta(25-35)$ or $A\beta(1-42)$ and $IFN-\gamma$ might mediate neuronal injury, we stimulated co-cultures of primary rat neuronal cells and N9 microglial cells with $A\beta$ peptides, or LPS, in the presence or absence of $IFN-\gamma$. Figure 4 shows neuronal cell injury observed by phase contrast microscopy after 72 h of incubation. Negligible signs of cell damage were observed when single activating signals were added to primary neurons alone (not shown) or in co-culture with microglia (Fig. 4a-e). In contrast, the combination of the various stimuli exerted a dramatic neurotoxic effect in the co-cultures, but not on the neurons, which was particularly marked with $A\beta$ s plus $IFN-\gamma$ (Fig. 4f, g), and maximal with LPS plus $IFN-\gamma$ (Fig. 4h). Neuronal cell loss was already evident 48 h after exposure to $A\beta$ s plus $IFN-\gamma$; in comparison, the LPS plus $IFN-\gamma$ combination had a very pronounced effect on neuronal mortality under the same conditions. Whereas $A\beta$ s plus $IFN-\gamma$ caused a great loss of neurites with fragmentation of cell bodies and the loss of neuronal cells, the combination of LPS

plus $IFN-\gamma$ induced the death of all neuronal cells, as previously shown¹¹⁻¹³, and even caused morphological signs of injury on the microglial cells themselves. To evaluate and quantify intermediate stages of neuronal cell injury, we also determined 3H -GABA (γ -aminobutyric acid) uptake, as an index of neuronal function. The reduction of 3H -GABA uptake in the co-cultures treated with $IFN-\gamma$ plus $A\beta(1-42)$, $A\beta(25-35)$ or LPS was marginal after 24 h (not shown), but by 48 h of incubation it was significantly reduced by, respectively, $23 \pm 6\%$ (mean $\pm$ s.e., $n = 4$), $20 \pm 4\%$ ($n = 5$) and $76 \pm 2\%$ ($n = 5$) (Fig. 4i). These results thus demonstrate that activation of microglia with $IFN-\gamma$ plus $A\beta$ may result in neuronal cell injury and degeneration.

We have identified novel biological effects exerted by $A\beta$ which may induce neuronal cell injury and degeneration through activation of microglia. Production of toxic free radicals by $A\beta$ -activated microglial cells supports the view that oxidative damage may contribute to the process of neuronal degeneration that occurs in normal ageing, Alzheimer's disease and similar condi-

tions such as Down's syndrome¹, and this agrees with other observations¹³⁻¹⁷. The production of TNF- α from microglia activated by A β observed in our study could be the source of TNF- α , and probably of other cytokines (IL-1 α/β and IL-6; refs 18, 19) present in senile plaques, but the role of TNF- α on neuronal behaviour remains controversial^{12,20}. In this regard, our data show that production of reactive nitrogen intermediates is in part mediated by the induced release of TNF- α . The observation that A β acted synergistically, or at least additively, with IFN- γ raises the problem of whether any normal products of neuronal catabolism other than IFN- γ may cooperate with A β in mediating its biological effects. □

Received 21 November 1994; accepted 21 February 1995.

1. Selkoe, D. J., Bell, D. S., Podlisny, M. B., Price, D. L. & Cork, L. C. *Science* **235**, 873-877 (1987).
2. Yankner, B. A. & Mesulam, M. M. *New Engl. J. Med.* **325**, 1849-1857 (1991).
3. Mullan, M. & Crawford, F. *Trends Neurosci.* **16**, 398-403 (1993).
4. Yankner, B. A., Duffy, L. K. & Kirschner, D. A. *Science* **250**, 279-282 (1990).
5. Kowall, N. W., Beal, M. F., Busciglio, J., Duffy, L. K. & Yankner, B. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7247-7251 (1991).
6. Pike, C. J., Burdick, D., Walenciewicz, A. J., Glabe, C. G. & Cotman, C. W. *J. Neurosci.* **13**, 1676-1687 (1993).
7. Perry, V. H., Andersson, P. B. & Gordon, S. *Trends Neurosci.* **16**, 268-273 (1993).

8. Corradini, S. B., Manuel, J., Donini, S. D., Quattrocchi, E. & Ricciardi-Castagnoli, P. *Glia* **7**, 255-262 (1993).
9. Simmons, L. K. et al. *Molec. Pharmacol.* **45**, 373-379 (1994).
10. Oswald, I. P., Wynn, T. A., Sher, A. & James, S. L. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8676-8680 (1992).
11. Chao, C. C., Hu, S., Molitor, T. W., Shaskan, E. G. & Peterson, P. K. *J. Immun.* **149**, 2736-2741 (1992).
12. Chao, C. C., Molitor, T. W. & Hu, S. *J. Immun.* **151**, 1473-1481 (1993).
13. Boje, K. M. & Arora, P. K. *Brain Res.* **587**, 250-256 (1992).
14. Olanow, C. W. *Trends Neurosci.* **16**, 439-444 (1993).
15. Coyle, J. T. & Puttfarcken, P. *Science* **262**, 689-692 (1993).
16. Hensley, K. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 3270-3273 (1994).
17. Behl, C., Davis, J. B., Lesley, R. & Schubert, D. *Cell* **77**, 817-827 (1994).
18. Aranjou, D. M. & Cotman, C. W. *Brain Res.* **569**, 141-145 (1992).
19. Dickson, D. W., Mattiace, L. A. & Yen, S. H. (abstr.) *Soc. Neurosci.* **16**, 1267 (1990).
20. Cheng, B., Christakos, S. & Mattson, M. P. *Neuron* **12**, 139-153 (1994).
21. Agresti, C., Aloisi, F. & Levi, G. *Dev. Biol.* **144**, 16-29 (1991).
22. Ding, A. H., Nathan, C. F. & Stuehr, D. J. *J. Immun.* **141**, 2407-2412 (1988).
23. Otvos, L., Szendrei, G. I., Lee, V. M. Y. & Mantsch, H. H. *Eur. J. Biochem.* **211**, 249-257 (1993).
24. Cassatella, M. A., Meda, L., Bonora, S., Ceska, M. & Constantin, G. *J. exp. Med.* **178**, 2207-2211 (1993).
25. Hernandez-Caselles, T. & Stutman, O. *J. Immun.* **151**, 3999-4012 (1993).
26. Villalba, M. et al. *J. Biol. Chem.* **269**, 2468-2476 (1994).
27. Sher, P. K. & Hu, S. *Brain Res. Bull.* **25**, 697-701 (1990).

ACKNOWLEDGEMENTS. This work was supported by grants from MURST, the CNR and the AIRC, by funds from Associazione Amici del Centro Dino Ferrari, and with the financial help of Alzheimer's Disease Research, a programme of the American Health Assistance Foundation. We thank P. Ricciardi-Castagnoli (Centro di CitoFarmacologia, Milano) for providing us with N9 and N13 microglia cell lines, F. Aloisi and R. Levi (Istituto Superiore Sanita, Roma) for primary murine and rat microglia cells, and G. Trinchieri (Wistar Institute, Philadelphia, USA) for his gift of anti-TNF- α monoclonal antibodies. We also thank P. P. McDonald for revising this manuscript.

Impairment of muscle function caused by mutations of phosphorylation sites in myosin regulatory light chain

Rutaiwan Tohtong*, Hiroshi Yamashita†, Melissa Graham*, Joe Haeberle†, Amanda Simcox* & David Maughan†‡

* Department of Molecular Genetics, The Ohio State University, Columbus, Ohio 43210, USA

† Department of Molecular Physiology and Biophysics, University of Vermont, Burlington, Vermont 05405, USA

MYOSIN regulatory light chain is phosphorylated by myosin light chain kinase at conserved serine and threonine residues in a number of species¹. Phosphorylation of myosin regulatory light chain regulates smooth muscle contraction, but appears to have a modulatory role in striated muscle contraction². We assessed the *in vivo* role of myosin regulatory light chain phosphorylation in the striated muscles of *Drosophila melanogaster* by substituting alanine at each or both conserved myosin light chain kinase-dependent phosphorylation sites, serine 66 and serine 67. We report here that myosin light chain kinase-dependent phosphorylation is not required for myofibrillogenesis or for the development of maximal isometric force in indirect flight muscles. However, mutants with substitutions at the major phosphorylation site (serine 66) or with the double substitutions had reduced power output in isolated flight muscle fibres and reduced flight ability, showing that myosin regulatory light chain phosphorylation is a key determinant of the stretch activation response in *Drosophila*.

In *Drosophila*, phosphorylation of the muscle-specific myosin regulatory light chain (called myosin light chain-2, or MLC2), has been correlated with a modest (1.5-2-fold) activation of myosin ATPase activity *in vitro*³ and acquisition of flight in young adult flies⁴. By sequence analysis, the consensus sites for myosin light chain kinase (MLCK) phosphorylation correspond to serine 66 (S66) and serine 67 (S67). Phosphopeptide mapping

of the bacterially expressed MLC2 shows both serines are phosphorylated by a Ca²⁺/calmodulin-dependent MLCK from rabbit skeletal muscle, with S66 phosphorylated to a greater extent than S67 (data not shown). We assessed the role of MLC2 phosphorylation *in vivo* by introducing mutant *Mlc2* genes, with serine to alanine substitutions, into a *Mlc2* null mutant⁵ (*Mlc2*^{E38}, hereafter called *Mlc2*⁻) using P-element mediated germline transformation. The *Mlc2*⁻ gene has a stop codon at residue 10 of the protein coding sequence⁵, and consequently in the transformants (*P[Mlc2*^{S66A}]; *Mlc2*⁻, *P[Mlc2*^{S67A}]; *Mlc2*⁻ and *P[Mlc2*^{S66A.S67A}]; *Mlc2*⁻, hereafter called *Mlc2*^{S66A}, *Mlc2*^{S67A} and *Mlc2*^{S66A.S67A}) the only full-length form of the protein produced is encoded by the introduced gene.

Mlc2⁻ homozygotes are embryonic lethal⁵. *Mlc2*^{-/+} heterozygotes are viable, but they are severely flight impaired and have a reduced wing beat frequency that is due to abnormal flight muscle structure⁵ (Fig. 1*b*; compare to wild type in Fig. 1*a*). Transforming *Mlc2*⁻ homozygotes with two wild-type (*Mlc2*⁺) genes rescued viability, restored wild-type muscle structure (not shown), and returned wing beat frequency (*f*_{wb}) and flight ability to normal (Table 1). *Mlc2*⁻ homozygotes transformed with two copies of any of the three mutant *Mlc2* genes were also viable, demonstrating that these serines are not critically important for muscles involved in normal development of the fly. As *Mlc2* is a single copy gene which encodes a single protein (see refs in ref. 6), another MLC2 isoform cannot mask the effect of the mutations in these muscles. However, multiple MLC2 phosphovariants (in our hands, at least 12; data not shown) exist in wild type⁶. In the mutant with the double substitution there was no consistent reduction in the number of phosphovariants but there was a more than threefold increase in the accumulation of the unphosphorylated variant (data not shown).

Indirect flight muscle (IFM) structure was normal in all three *Mlc2* mutants (for example *Mlc2*^{S66A.S67A}; Fig. 1*c*), indicating that the conserved serines are not important for myofibrillogenesis. This normal muscle structure has allowed us to interpret functional phenotypes in the mutants as a specific consequence of the lack of phosphorylation rather than a general effect due to structural disruption. Indeed, we found significant mutant phenotypes. The mutant with the double substitution (*Mlc2*^{S66A.S67A}) was flightless (Table 1). To our knowledge, this is the first example of a flightless phenotype in a fly with structurally normal IFM (a near normal myofibrillar structure has been reported for a flightless actin mutant⁶). We also analysed flight

‡ To whom correspondence should be addressed.

TABLE 1 Flight ability, wing beat and isolated flight muscle mechanical performance of *Drosophila* transformed with wild-type or mutant myosin regulatory light chain genes (means $\pm$ s.e.m.)

Tests of	Whole fly			Isolated flight muscle fibre*				
	Flight index		Wing beat frequency† (Hz)	f_0 (Hz)	$E_v(f_0)$ (kN m ⁻²)	f_{max} (Hz)	P_{max} (W m ⁻³)	E_{max} (kN m ⁻²)
Functional assay	Group‡	Individual§						
Strain								
<i>Mlc2</i> ⁺	5.6 $\pm$ 0.5	5.4 $\pm$ 0.8	208 $\pm$ 16	41 $\pm$ 3	-92 $\pm$ 16	56 $\pm$ 3	22 $\pm$ 4	476 $\pm$ 34
<i>Mlc2</i> ^{S67A}	6.8 $\pm$ 0.6	5.4 $\pm$ 0.9	215 $\pm$ 10	53 $\pm$ 6	-58 $\pm$ 9	70 $\pm$ 6	17 $\pm$ 2	388 $\pm$ 50
<i>Mlc2</i> ^{S66A}	3.2 $\pm$ 0.6**	3.0 $\pm$ 1.0**	192 $\pm$ 7**	43 $\pm$ 5	-37 $\pm$ 7**	57 $\pm$ 5	11 $\pm$ 2**	390 $\pm$ 54
<i>Mlc2</i> ^{S66A,S67A}	0.2 $\pm$ 0.2**	0.9 $\pm$ 0.6**	173 $\pm$ 11**	42 $\pm$ 4	-11 $\pm$ 2**	42 $\pm$ 4**	2 $\pm$ 0.4**	363 $\pm$ 50

Elastic and viscous stiffnesses were calculated at 31.6 μ M Ca²⁺ by computing the amplitude ratio and the phase difference for force and length at each frequency. Stiffnesses were converted to moduli by dividing the amplitude ratio by fibre cross-sectional area (from width measurements in oil) and multiplying the amplitude ratio by initial (zero stress) fibre length. The elastic and viscous moduli are the real and imaginary part, respectively, of the amplitude ratio multiplied by the exponential phase difference²⁶. f_0 is the frequency at which the negative viscous modulus is maximum. $E_v(f_0)$ is the viscous modulus at f_0 , indicated by the dotted line in Fig. 2. Power (in watts) = $2\pi f E_v A L ((\Delta L/L)_{r.m.s.})^2$, where f is the frequency of the length perturbation (in s⁻¹), A is the fibre cross-sectional area (in m²), L is the fibre length between T-clips (in m), and $\Delta L/L$ is the amplitude of the sinusoidal length perturbation divided by L (ref. 17). r.m.s., Root-mean-square of the perturbation amplitude; P_{max} , power output (in watts m⁻³), normalized with respect to A and L at the frequency at which power is maximum (f_{max}). The dynamic stiffness modulus (E_{max}) is the vector sum of the elastic and viscous moduli at f_{max} . The elastic modulus at f_0 in all three mutant strains was not significantly different from that of *Mlc2*⁺ (508 $\pm$ 39 kN m⁻²). See refs 5, 14, 17 and 26 for further details of the method of sinusoidal analysis. Note that the results obtained from the isolated fibres are similar to those of the flight and wing beat tests, and are consistent with a previous report that *Drosophila* cannot fly given only half its normal power output²⁹. Also note f_{wb} is roughly proportional to the square root of the active dynamic stiffness modulus of the IFM (compare column 3 and 8), consistent with active dynamic stiffness being the primary contributor to flight system stiffness and a major determinant of f_{wb} ³⁰.

* Results of 6–7 single fibres from each strain.

† For wing beat analysis, each fly was tethered by a leg and the wing movement recorded by an optical device; the wing beat frequency (f_{wb}) was extracted using a spectrum analysis routine²⁸.

‡ About 100 flies tested, released in groups of 2 to 3 inside the top of a graduated cylinder and scored according to where each landed (8 = top, 0 = bottom); the index listed is the average score of all flies tested for the representative line of each transformant type. Altogether, flight indices were determined for 4 lines of *Mlc2*⁺, 3 lines of *Mlc2*^{S66A}, 6 lines of *Mlc2*^{S67A} and 8 lines of *Mlc2*^{S66A,S67A}; for each transformant strain, all lines exhibited essentially the same group flight index, with slight variations due to effects arising from different chromosomal insertion sites. Lines with the highest group indices were used for the individual flight tests and isolated fibre experiments (columns 2, 4–8).

§ Score from repeated trials for each fly; the index is the grand average of scores from 10–40 individuals, each scored according to the formula: $6 U/T + 4 H/T + 2 D/T + 0 N/T$, where each individual either flew up (U), horizontally (H), down (D), or not at all (N)²⁷. T is the total number of trials done on each fly.

|| f_{wb} of flies exhibiting wing beats; about 1 in 10 mutant flies with the double substitution could not be induced to beat their wings. Temperature columns 1–3, 22 °C; columns 4–8, 12 °C.

** $P < 0.05$ versus control (*Mlc2*⁺ homozygote).

in the mutants with single substitutions (Table 1). *Mlc2*^{S66A} exhibited a significant reduction in f_{wb} and flight ability, although the dysfunction was not as severe as that of *Mlc2*^{S66A,S67A}. In contrast, *Mlc2*^{S67A} displayed a f_{wb} and flight ability that were not significantly different from those of *Mlc2*⁺. Thus, of the two serines, S66 appears to have greater functional importance.

To determine the direct effects of the MLC2 mutations on flight muscle performance, we measured both the Ca²⁺-dependent isometric tension and the power output associated with oscillatory (sinusoidal) length changes of demembrated IFM fibres at 12 °C.

Ca²⁺-activated isometric tension increased sigmoidally from zero tension at $\sim 0.1 \mu$ M Ca²⁺ to a maximal response of

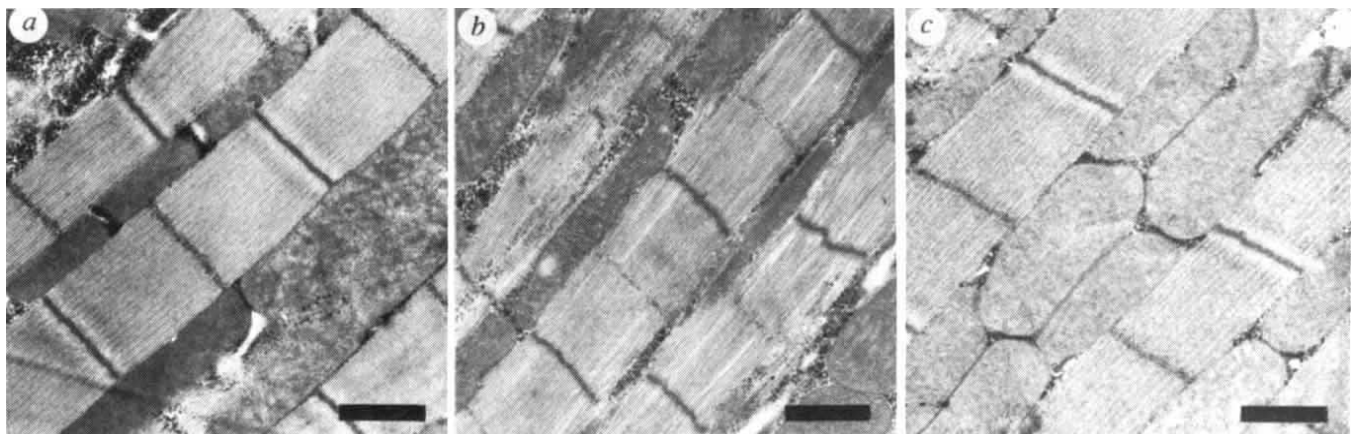
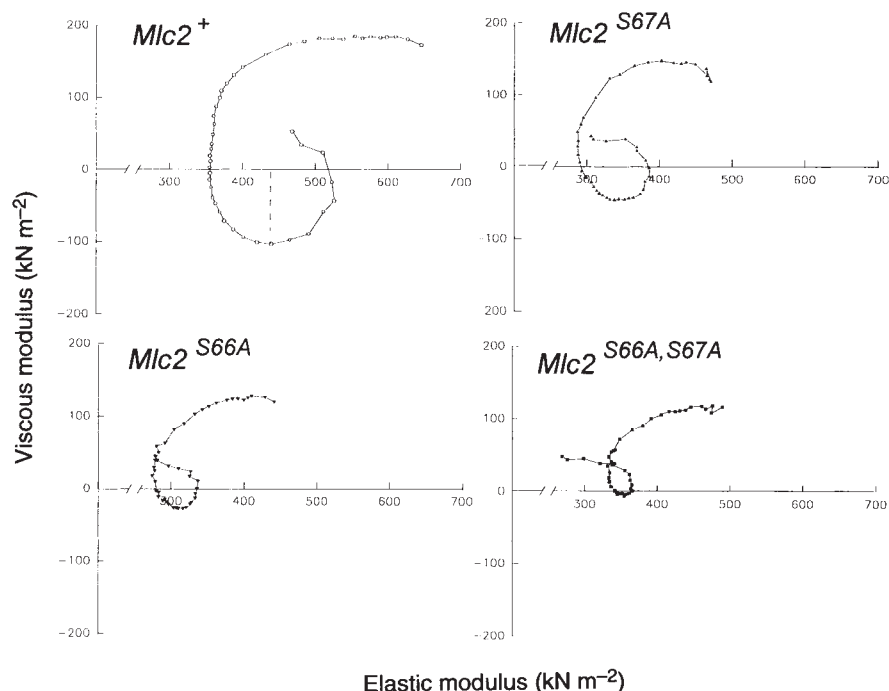


FIG. 1 Electron micrographs showing longitudinal sections of indirect flight muscles from wild-type (Canton S) (a), *Mlc2*^{-/-} heterozygote (b) and *Mlc2*^{S66A,S67A} homozygote (c). The wild-type structure (a) appears well organized with compact arrays of interdigitating thick and thin filaments running parallel to the longitudinal axis of the myofibrils. The Z bands and the M lines are well defined and straight, spanning the entire width of the myofibrils. In contrast, *Mlc2*^{-/-} myofibrils (b) are less compact, and gaps between filament arrays are often seen. Core regions of the myofibrils are well ordered, but this organization is lost in the

periphery. Z bands and M lines sometimes appear wavy or do not span the entire width of the myofibrils. Broken or branched Z bands are also common. The structure of the *Mlc2*^{S66A,S67A} fibres (c) is similar to that of the wild-type (a). The organization of the thick and thin filaments within the myofibrils as well as the appearance of the Z bands and the M lines are indistinguishable from those of *Mlc2*⁺ transformants with two copies of the wild-type genes (not shown) and Canton S (a). Sample preparation for electron microscopy was described previously⁵. Scale bar, 1 μ m.

FIG. 2 Sinusoidal analysis of single demembrated cells from the indirect flight muscle of *Drosophila* at 12 °C. The homozygous transformants contained either two copies of the wild-type myosin regulatory light chain gene (*Mlc2*⁺) or two copies of the mutant *Mlc2* genes (*Mlc2*^{S66A}, *Mlc2*^{S67A} or *Mlc2*^{S66A,S67A}). The genetic background (host strain) was the *Mlc2* null mutant *Mlc2*^{E38} (ref. 5). The amplitude of the viscous modulus at which work output is maximum ($-E_v(f_o)$, Table 1) is indicated by the dotted line in *Mlc2*⁺. Power output of *Mlc2*⁺, calculated from the product of E_v and frequency (Table 1 caption), was ~22 W m⁻³. This value is like that measured in wild-type (Canton S) IFM under similar conditions (data not shown), and appears to be consistent with other values reported¹⁴ if effects due to differences in ionic strength and perturbation amplitudes are taken into account.

METHODS. Fibres were chemically demembrated (skinned) for 2 h (10 °C) in a relaxing solution containing ~0.01 μ M Ca²⁺, 100 μ g ml⁻¹ saponin, 5 mM Mg-ATP, 15 mM phosphocreatine, 80 units ml⁻¹ creatine phosphokinase, 5 mM EGTA, 10 mM DTT, 0.2 mM leupeptin, 50% glycerol and 20 mM BES (*N,N*-bis(2-hydroxyethyl)-2-aminoethanesulphonic acid) buffer, pH 7.0, 175 mM ionic strength adjusted with K methane sulphate. The skinned fibres, clamped between aluminium T-clips, were mounted on a mechanical rig similar to that described in ref. 5, and stretched taut (~15%) to a level of resting stress equivalent to ~1 kN m⁻² after stress relaxation. The isometric and dynamic force responses of the fibres were then analysed in a series of graded activating solutions (~0.01 μ M–31.6 μ M Ca²⁺), culminating at 31.6 μ M Ca²⁺. Activating solutions contained 16 mM free phosphate in addition to the above constituents, but lacked saponin, leupeptin and glycerol; CaCl₂



2.0 ± 0.3 kN m⁻² (*Mlc2*⁺ control fibres) at ~10 μ M Ca²⁺, with no significant difference ($P > 0.05$) in response between control and the *Mlc2* mutants. This insensitivity of Ca²⁺-activated tension to MLC2 phosphorylation site mutation is reminiscent of results obtained from vertebrate striated muscle, where MLC2 phosphorylation has no effect on isometric tension at maximal Ca²⁺ activation^{2,7,10} and a very slight enhancement of tension at submaximal Ca²⁺ activation at low temperature (15 °C)².

In isolated Ca²⁺-activated IFM, as in vertebrate striated muscles, the response to a rapid stretch consists of a tension increase concurrent with the stretch, followed by a rapid decay of tension and then a delayed, sustained rise in tension^{11,15} (generally called 'stretch activation'¹⁶). The delayed tension underlies the ability of all striated muscles to do oscillatory work. In insect IFM stretch activation is especially prominent¹¹ and is required for flight. We measured the stiffness moduli and determined the power output (work per unit time) of Ca²⁺-activated IFM during sinusoidal length perturbations applied at various frequencies^{12,14,17}. Nyquist plots from the sinusoidal analysis^{5,12,14,17} showed a graded effect of the *Mlc2* mutations on the amplitude of the viscous stiffness modulus (E_v) (Fig. 2), with the largest attenuation of the maximum negative value in *Mlc2*^{S66A,S67A} and the least in *Mlc2*^{S67A} (Table 1, column 5). The graded reduction in $-E_v$ corresponds to a graded reduction in the amplitude of tension generated in response to stretch (data not shown) and power output (Table 1, column 7). The reduced $-E_v$ implies that during each wing beat in the fly less work will be done, causing the flight phenotypes seen in the mutants.

The frequency at which $-E_v$ was maximum (f_o) was not affected significantly by the mutations (Table 1, column 4), suggesting that the phosphorylation site mutations affect primarily the amplitude of the tension response to stretch (proportional to $-E_v$) rather than the kinetics of the response (proportional to f_o)^{12,14}. These data suggest that stretch activation results from

recruitment of crossbridges with unchanged kinetic properties¹⁵. The important result of this study is that mutation of specific phosphorylatable serine residues causes a large reduction in the stretch activation response of the IFM without changing isometric tension. Although it is unclear why this response is so pronounced in IFM, it is possible that thick filament strain, transmitted through connecting filaments, induces matching of the actin and myosin filament helical structure (the 'Wray' model¹⁸), thereby increasing the probability of crossbridge attachment (refs 12, 13, 18, 19, but see 20, 21). Alternatively, thick filament strain may directly enhance the rate at which force-generating crossbridges are formed^{12,13}. A third possibility is that linkage of the thick and thin filaments in IFM may amplify the response to strain. These links may consist of IFM-specific thin filament-associated proteins and a thick filament component which we propose is phosphorylated MLC2. This linkage could function as a strain-sensitive element^{22–24}. It may also contribute to stretch activation by increasing the binding of myosin to actin by maintaining the myosin head close to the actin filament because crosslinking as little as 18% of the myosin heads to the thin filament produces an insect-like stretch activation response in vertebrate striated muscle²⁵.

We favour the third type of model and propose that MLC2 phosphorylation is required for forming the inter-filament links. However, our data do not exclude the other models or a mechanism whereby MLC2 phosphorylation could also affect the mobility or position of the myosin head to promote actomyosin interaction^{2,10}. *Drosophila* MLC2 has an N-terminal extension⁶ not conserved in the vertebrate proteins which may be important for the inter-filament link. The unique N-terminal extension has the potential to interact electrostatically and hydrophobically with the region around the MLCK substrate serines, promoting a folded conformation. Phosphorylation of serine residues could relieve the intramolecular interaction and allow the N terminus

to extend. Much of the N-terminal extension is hydrophobic and thus could associate with similar domains in the IFM-specific thin filament proteins (troponin I and/or tropomyosin isoforms⁶ by hydrophobic interaction). This postulated link between the thick and thin filaments may account for the pronounced stretch activation response of insect flight muscle^{20,23,24} and its functional dependence on MLC2 phosphorylation in *Drosophila*. □

Received 16 January; accepted 13 February 1995.

- Collins, J. H. *J. Muscle Res. Cell Motil.* **12**, 3–25 (1991).
- Sweeney, H. L., Bowman, B. F. & Stull, J. T. *Am. J. Physiol.* **264**, C1085–C1095 (1993).
- Takahashi, S., Takano-Ohmuro, H. & Maruyama, K. *Comp. Biochem. Physiol.* **95B**, 179–181 (1990).
- Takahashi, S. & Maruyama, K. *Zool. Sci.* **4**, 833–838 (1987).
- Warmke, J., Yamakawa, M., Molloy, J. E., Falkenthal, S. & Maughan, D. W. *J. Cell Biol.* **119**, 1523–1536 (1992).
- Bernstein, S. I., O'Donnell, P. T. & Cripps, R. M. *Int. Rev. Cytol.* **143**, 63–152 (1993).
- Persechini, A., Stull, J. T. & Cooke, R. J. *biochem. J.* **260**, 7951–7954 (1985).
- Sweeney, H. L. & Stull, J. T. *Am. J. Physiol.* **250**, C657–C660 (1986).
- Sweeney, H. L. & Stull, J. T. *Proc. natn. Acad. Sci. U.S.A.* **87**, 414–418 (1990).
- Metzger, C. J., Greaser, M. L. & Moss, R. L. *J. gen. Physiol.* **93**, 855–883 (1989).
- Steiger, C. J. in *Insect Flight Muscle* (ed. Tregear, R. T.) 221–268 (North-Holland, Amsterdam, 1977).

- Thorson, J. & White, D. C. S. *J. Physiol.* **343**, 59–84 (1983).
- Thorson, J. & White, D. C. S. *Biophys. J.* **9**, 360–390 (1969).
- Peckham, M., Molloy, J. E., Sparrow, J. C. & White, D. C. S. *J. Muscle Res. Cell Motil.* **11**, 203–215 (1990).
- Guth, K., Kuhn, H. J., Tsuchiya, T. & Ruegg, J. C. *Biophys. struct. Mech.* **7**, 139–169 (1981).
- Pringle, J. W. S. *Proc. R. Soc. B* **201**, 107–130 (1978).
- Jewell, B. R. & Ruegg, J. C. *Proc. R. Soc. B* **164**, 428–459 (1966).
- Wray, J. S. *Nature* **280**, 325–326 (1979).
- Abbott, R. H. & Cage, P. E. *J. Muscle Res. Cell Motil.* **5**, 387–397 (1984).
- Grazier, H. L. M. & Wang, K. J. *gen. Physiol.* **101**, 235–270 (1993).
- Squire, J. M. *J. Muscle Res. Cell Motil.* **13**, 183–189 (1992).
- Bullard, B. et al. *J. molec. Biol.* **204**, 621–637 (1988).
- Peckham, M., Cripps, R., White, D. & Bullard, B. J. *exp. Biol.* **168**, 57–76 (1992).
- Reedy, M. C., Reedy, M. K., Leonard, K. R. & Bullard, B. J. *molec. Biol.* **239**, 52–67 (1994).
- Tawada, K. & Kawai, M. *Biophys. J.* **57**, 643–647 (1990).
- Kawai, M. & Brandt, P. W. *J. Muscle Res. Cell Motil.* **1**, 279–303 (1980).
- Drummond, D. R., Hennessey, E. S. & Sparrow, J. C. *Molec. gen. Genet.* **226**, 70–80 (1991).
- Hyatt, C. J. & Maughan, D. W. *Biophys. J.* **67**, 1149–1154 (1994).
- Laurie-Ahlberg, C. C. et al. *Genetics* **111**, 845–868 (1985).
- Molloy, J., Kreuz, A., Miller, R., Tansey, T. & Maughan, D. in *Mechanism of Myofibrillar Sliding in Muscle Contraction* (eds Sugi, H. & Pollack, G. H.) 165–172 (Plenum, New York, 1993).

ACKNOWLEDGEMENTS. We thank C. Beall, K. Bickford, S. Brown, E. Fyrberg, M. Glackin, C. Hyatt, J. Hurley, M. Kawai, B. Oakley, M. Reedy, M. Reedy and J. Vigoreaux for their assistance. We also thank M. Chun and J. Warmke for providing the electron micrographs of Canton S and *Mic2^{+/+}*, respectively, and J. Sellers for the rabbit MLCK. This paper is dedicated to the late S. Falkenthal, who helped initiate this work. Supported by the NIH.

A TBP–TAF complex required for transcription of human snRNA genes by RNA polymerases II and III

R. William Henry*, Cynthia L. Sadowski†, Ryuji Kobayashi* & Nouria Hernandez‡

* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

† Department of Pharmacology, State University of New York at Stony Brook, Stony Brook, New York 11794, USA

‡ Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

THE TATA-box-binding protein TBP exists in the cell complexed with different sets of TBP-associated factors (TAFs)¹. In general, each of these TBP–TAF complexes is dedicated to transcription by a single RNA polymerase. Thus, SL1, TFIID and TFIIB are required for transcription by polymerases I, II and III, respectively^{2–10}. Here we characterize a fourth TBP–TAF complex called SNAPc¹¹. Unlike the other TBP–TAF complexes, SNAPc is implicated in transcription by two types of polymerases; it is required for transcription of both the RNA polymerase II and III small-nuclear RNA genes and binds specifically to the proximal sequence element PSE, a non-TATA-box basal promoter element common to these two types of genes^{11–13}. In addition to TBP, SNAPc is composed of at least three TAFs, SNAP43, SNAP45 and SNAP50. The predicted amino-acid sequence of SNAP43 reveals that it corresponds to a new protein.

We used conventional chromatographic methods to purify the small-nuclear (sn) RNA-activating protein complex (SNAPc) extensively from HeLa cell extracts. Figure 1a shows our purification scheme and Fig. 1b–f shows the analysis of fractions from a glycerol gradient, the last purification step. PSE binding activity peaked in fractions 16 and 17 (Fig. 1b), as did reconstitution of U1 transcription by RNA polymerase II (Fig. 1c). Immunoblot analysis revealed that fractions 16 and 17 also contained the peak of TBP (Fig. 1d), consistent with TBP being a component of the SNAP complex. To identify other potential components of SNAPc, the glycerol gradient fractions were analysed by SDS–PAGE and silver staining. Five proteins with apparent relative molecular masses (M_r) of 200K, 54K, 50K, 45K and 43K peaked in fractions 16 and 17 (bands labelled with dots in

Fig. 1e). Figure 1f shows the most prominent polypeptides in fractions 16 and 17, p43, p45 and p50.

In our gels, TBP migrates with an apparent M_r of 45K (Fig. 3 legend), but peptide sequence analysis of the 45K band yielded amino-acid sequences unrelated to TBP (unpublished results), indicating that p45 is a mixture of at least two polypeptides. The protein with an apparent M_r of 54K (faintly visible in Fig. 1e) is related to the 50K protein, because some identical peptide sequences were obtained from both of these proteins (unpublished results). Its abundance is variable in different SNAPc preparations, suggesting that it corresponds to a modified form of the 50K protein. Alternatively, the 50K protein may correspond to a degraded form of the 54K protein. The 200K protein (very faint band in Fig. 1e) also peaks in fractions 16 and 17 and may therefore be part of the complex.

We sequenced p43 and used this information to select two λ phage recombinants containing corresponding cDNAs. DNA sequence analysis revealed a single open reading frame encoding a novel protein of 368 amino acids with a predicted M_r of 42.9K (Fig. 2). The sequence contains all the p43 peptides we sequenced (underlined in Fig. 2), suggesting that the p43 band contains only one polypeptide. The carboxy-terminal two-thirds of p43 are highly charged overall, with 40% of the last 268 amino acids charged; this number increases to 47% for the last 64 amino acids.

To show that the 43K protein is part of the SNAP complex, we raised rabbit polyclonal antibodies against a synthetic peptide, CSH375, derived from the C-terminal region of the p43 amino-acid sequence (underlined with a shaded box in Fig. 2). The anti-CSH375 antibody supershifted the SNAPc–PSE complex in a gel mobility-shift assay and the supershift was inhibited by preincubation of the antibodies with the peptide antigen but not with a peptide derived from the central region of the p43 amino-acid sequence (CSH374; underlined with a shaded box in Fig. 2), indicating that the reaction was specific and that p43 is indeed part of the SNAP complex (data not shown). We therefore refer to the p43 protein as SNAP43.

We verified that the anti-CSH375 antibodies do not crossreact with TBP (data not shown) and used them to examine whether free SNAPc contains TBP. Indeed, SNAPc was originally suggested to be a TBP-containing multisubunit complex because when bound to the PSE in an electrophoretic mobility-shift assay (EMSA), it could be disrupted into faster-migrating partial complexes by addition of anti-TBP monoclonal antibodies¹¹. But, because the anti-TBP antibodies, unlike anti-CSH375 antibodies, disrupt rather than supershift the complex, we were unable to use them to immunopurify SNAPc and show that the complex contains TBP in the absence of the PSE binding site.

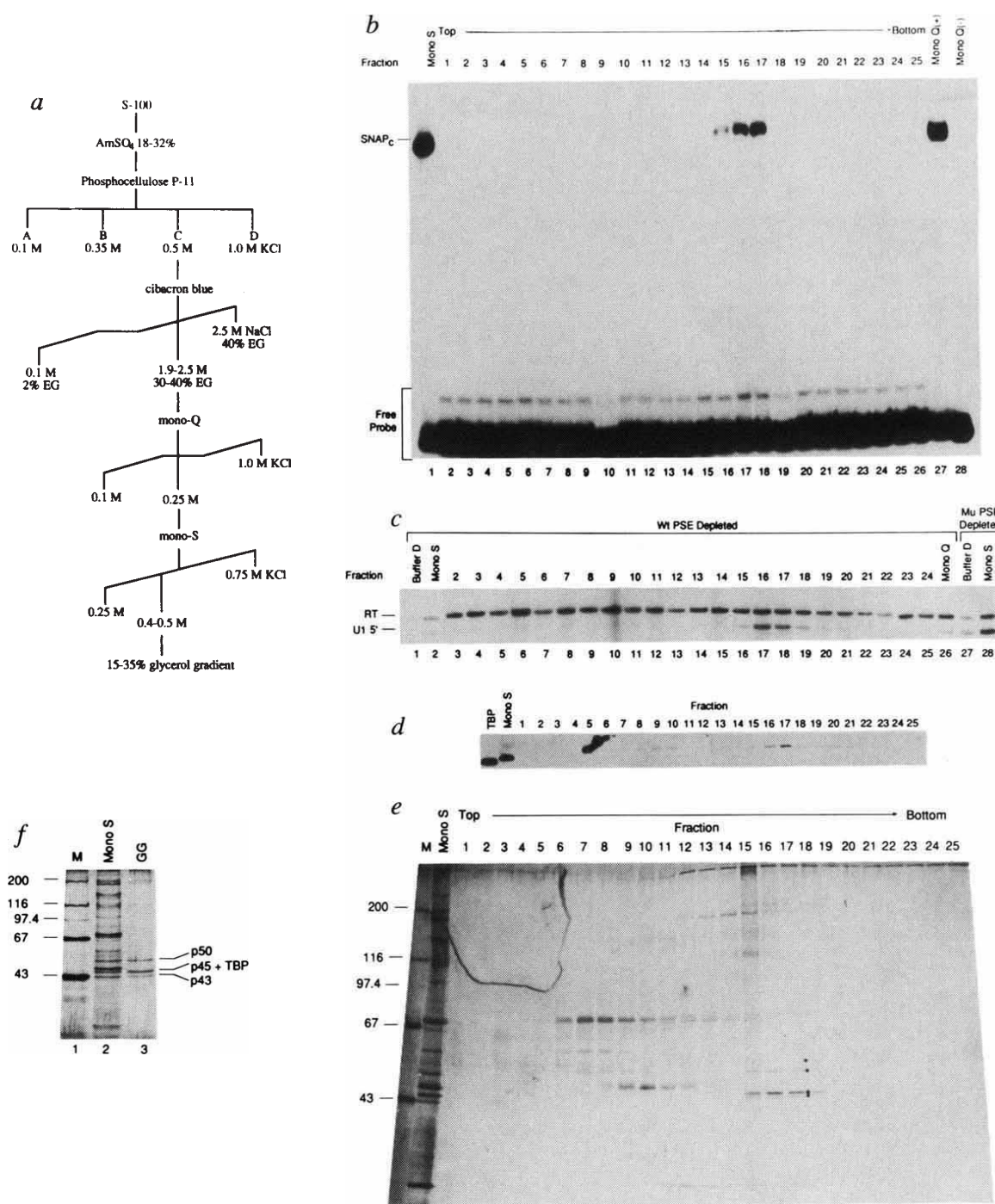


FIG. 1 *a*, SNAPc purification scheme. *b*, PSE binding activity peaks in glycerol gradient fractions 16 and 17. The peak mono-S fraction (lane 1) and each of the glycerol gradient fractions (lanes 2–26) were tested by EMSA with the wild-type PSE probe as described¹¹. In lanes 27 and 28, mono-Q peak fractions were assayed with a wild-type and mutant PSE probe, respectively. *c*, U1 transcriptional activity peaks in fractions 16 and 17. HeLa cell nuclear extracts depleted with either multimerized wild-type (lanes 1–26) or mutant (lanes 27, 28) PSE oligonucleotides crosslinked to Sepharose-4B beads were complemented with buffer (lanes 1 and 27), the mono-S (lanes 2 and 28) or mono-Q (lane 26) peak fractions, or the glycerol gradient fractions (lanes 3–25), and assayed for U1 transcription¹¹. 'RT' corresponds to RNAs initiated at cryptic promoters within vector sequences; 'U1 5'' corresponds to RNA correctly initiated at the U1 promoter¹¹. *d*, TBP peaks in fractions 16 and 17. 2 ng recombinant TBP (lane labelled TBP), 50 μ l of the peak mono-S fraction (lane labelled mono S) and 250 μ l of the glycerol gradient fractions (lanes labelled 1–25) were analysed by immunoblot with an anti-TBP mAb (SL39) by Enhanced ChemiLuminescence (Amersham). Note that the recombinant TBP used in the first lane is slightly shorter than wild-type TBP in the N-terminal region and therefore migrates faster on the gel. *e*, SDS-PAGE

analysis of glycerol gradient fractions. The mono-S peak fraction and the glycerol gradient fractions were fractionated by 12.5% SDS-PAGE and the proteins were stained with silver. *f*, The human SNAPc complex is composed of at least four polypeptides. The peak mono-S (lane 2) and glycerol gradient (fractions 16 and 17, lane 3) fractions were fractionated by SDS-PAGE and the proteins were stained with silver.

METHODS. Purification of SNAPc. SNAPc was purified as described¹¹ with the following modifications: (1) the cibacron blue affi-gel column (Sigma) was developed with a linear gradient in 20 mM HEPES, pH 7.9, 0.1% Tween 20, 0.2 mM EDTA from 2% ethylene glycol (v/v), 0.1 M NaCl to 40% ethylene glycol, 2.5 M NaCl, with a 5-column volume wash at 15% ethylene glycol, 0.95 M NaCl. SNAPc eluted between 1.9 and 2.4 M NaCl; (2) the mono-Q column was developed with a linear gradient in Q buffer from 100 mM KCl to 1 M KCl interrupted by a 5-column volume wash at 250 mM KCl; SNAPc eluted in the 250 mM wash. This fraction was loaded directly onto a mono-S column (Pharmacia) and eluted with a linear gradient from 250–750 mM KCl in Q buffer. The peak of SNAPc activity eluted between 350 and 450 mM KCl. This was then loaded onto a 15–35% glycerol gradient and centrifuged for 21 h at 49,000 r.p.m. in a SW55Ti rotor (Beckman).

CGGAGCGTGC GGCGCTTCGGGTGCC

ATGGGACTCTCCCGCCTGCAGACCGACTGCGAGGCGCTGCTCAGCCGCTTCCAGGAG
M G T P P G L Q T D C E A L L S R F Q E 20

ACGGACAGTGTACGCTTCAGGAGCTTCACGGAGCTCTGGAGAAACATGAAGTTCGGGACT
T D S V R F E D F T E L W R N M K F G T 40

ATCTTCTGGCAGAATGAGAAATTTAGAAAAGAACATGTTTACAAAAGAGCTTTAGCT
I F C G R M R N L E K N M F T K E A L A 60

TTGGCTTGGCGATATTTTACCTCCATACACCTTCCAGATCAGAGTTGGTGTCTTGTAT
L A W R Y F L P P Y T F Q I R V G A L Y 80

CTGCTATATGATTATATAATACCACTGTGTCAACCAAAACAAAGATCAGAGTTGCC
L L Y G L Y N T Q L C Q P K Q K I R V A 100

CTGAAGGATGGGATGAAGTTTAAATTTTCAGCAAGATTTAGTAAATGCACAGCATTTT
L K D W D E V L K F Q Q D L V N A Q H F 120

GATGCAGCTTATATTTTAGGAAGCTACGACTAGACAGAGCATTTTCACTTTACAGCAATG
D A A Y I F R K L R L D R A F H F T A M 140

CCCAATTTGCTGTCATATAGGATGAAGAAAAATTCACCGAGCTGAAGTTACAGAAGAA
P K L L S Y R M K K K I H R A E V T E E 160

TTTAAAGACCAAGTGATCGTGTGATGAACCTTATCACTTCTGATGATTAGAGGAAATG
F K D P S D R V M K L I T S D V L E E M 180

CTGAATGTTTCATGATCATTATCAGAACATGAAACATGTAATTTTCAGTTGTAAGTCCAAG
L N V H D H Y Q N M K H V I S V D K S K 200

CCAGATAAAGCCCTCAGCTTGATAAAGGATGATTTTGTGACATATTAAGACATAGTT
P D K A L S L I K D D F D N I K N I V 220

TTGGAGCATCAGCAGTGGCACAAGACAGAAAGAAATCCATCCTTAAAGTCAAAACTAAT
L E H Q Q W H K D R K N P S L K S K T N 240

GATGGAGAAGAAAAATGGAAGAAATTCACAAGAAACGGAGAGATGTGAAAGGGCAGAA
D G E E K M E G N S Q E T E R C E R A E 260

TCATTAGCGAAAATAAATCAAGGCCCTTTTCAGTTGTCATACAGGCATCCAAATCAAGA
S L A K I K S K A F S V V I Q A S K S R 280

AGGCATCGTCAAGTCAAATCGACTCTTCTGACTCTGATTCGATCTGGTCAAGGGCAA
R H R Q V K L D S S D S A S G Q G Q 300

GTCAAAGCAACTAGGAAAAAGAGAAGAAAGAAATGAAACCAAGCAGGAAGGAAGATG
V K A T R K K E K K E R L K P A G R K M 320

TCCTCTCAAGAAACAAAGGCAATGTGCAGAAATATACACAAGGAAGATAAACCTTTAAGTCTG
S L R N K G N V Q N I H K E D K P L S L 340

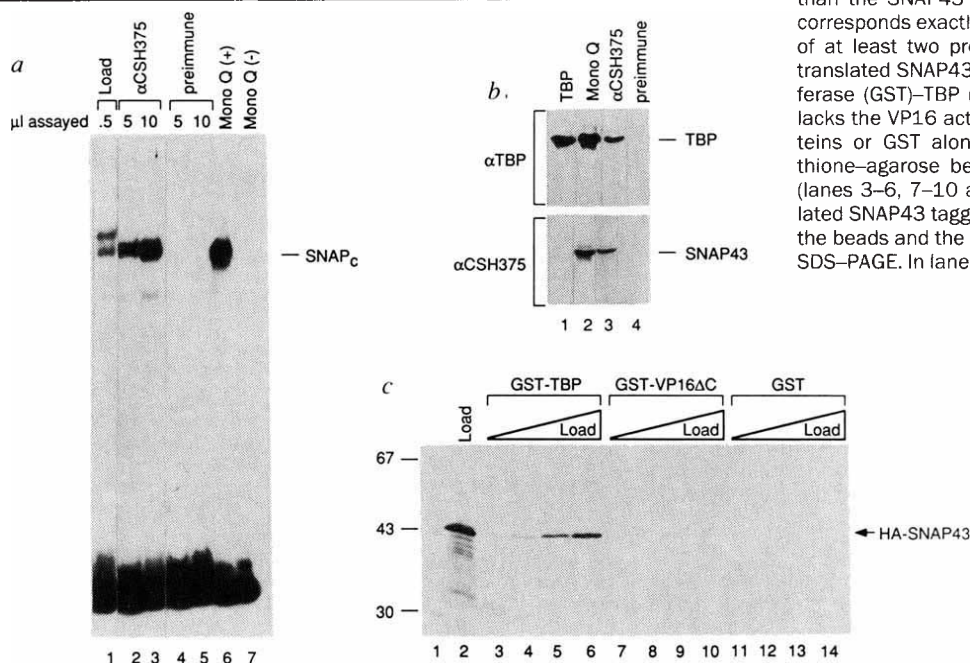
AGTATGCCTGTAATTACAGAAGAAGAGAATGAAAGTTTGAGTGAACAGAGTTCACT
S M P V I T E E E E N E S L S G T E F T 360

GCATCCAAGAAGAGGAGAAAACACTGA
A S K K R R K H * 368

FIG. 2 Nucleic acid sequence and predicted amino-acid sequence of SNAP43. The peptide sequences obtained from the purified protein are underlined. The sequences corresponding to the two synthetic peptides CSH374 and CSH375 are underlined with shaded boxes.

METHODS. The mono-S peak of PSE binding activity was loaded on a 12.5% SDS-polyacrylamide gel, stained with 0.05% Coomassie blue G for 15 min, destained and soaked in water for 1 h. The bands corresponding to bands marked with a dot in Fig. 1e were excised from the gel and subjected to digestion *in situ* with lysylendopeptidase. The resulting peptides were separated by reverse-phase HPLC and sequenced with an ABI automated protein sequencer as described¹⁴. Degenerate oligonucleotides were designed from the amino-acid sequences of the purified p43 polypeptide. The oligonucleotides were used for polymerase chain reaction (PCR) with cDNA generated from total RNA as a template. A specific 231-bp fragment (corresponding to the sequence encoding amino acids 152–228) was obtained and used as a probe to screen a human cDNA library from Ntera2D1 cells¹⁵, a human teratocarcinoma cell line. Out of 750,000 plaques screened, 14 were positive, and the two that gave the strongest hybridization signal were analysed further. These two positive λ recombinants contained inserts of different sizes which were subcloned into pUC118 for nucleic acid sequencing.

FIG. 3 TBP is part of the SNAP complex. **a**, Immunopurification of SNAPc from nuclear extract. 10 ml HeLa cell nuclear extract were mixed with 50 μ l protein-A-agarose beads (Boehringer Mannheim) crosslinked to anti-CSH375 (α -CSH375, lanes 2, 3) or preimmune (lanes 4, 5) antibodies. After extensive washing, the beads were eluted with 250 μ l buffer containing 1 mg ml⁻¹ of the specific peptide CSH375, and the eluate tested in an EMSA. Nuclear extract (lane 1) and the mono-Q peak fraction (lanes 6, 7) were also tested. In lanes 1–6 and 7, wild-type and mutant PSE probes were used, respectively. **b**, Anti-TBP and anti-CSH375 immunoblot analysis of immunopurified SNAPc. Full-length recombinant TBP (3 ng; Promega) (lane 1), 50 μ l peak mono-Q fraction (lane 2), and 235 μ l material eluted from the anti-CSH375 (lane 3) or control (lane 4) beads were fractionated by SDS-PAGE, the proteins were transferred to nitrocellulose, and the filter was probed first with a mixture of two anti-TBP mAbs (SL20 and SL39) (top), stripped and reprobed with the anti-CSH375 antibodies purified by chromatography over protein-A (bottom). Note that owing to incomplete stripping, TBP is still visible in the second panel as a very weak band of slightly slower mobility than the SNAP43 band. In silver-stained gels, this signal corresponds exactly to the p45 band. Thus, p45 is a mixture of at least two proteins, one of which is TBP. **c**, *In vitro* translated SNAP43 interacts with TBP. Glutathione-S-transferase (GST)-TBP (lanes 3–6) and GST-VP16 Δ C (VP16 Δ C lacks the VP16 activation domain) (lanes 7–10) fusion proteins or GST alone (lanes 11–14) were linked to glutathione-agarose beads (Sigma). Then 1, 3, 10 and 30 μ l (lanes 3–6, 7–10 and 11–14) of radioactive *in vitro* translated SNAP43 tagged with an HA epitope¹⁶ were mixed with the beads and the bound material was fractionated by 15% SDS-PAGE. In lane 2, loading was 10 μ l translated material.



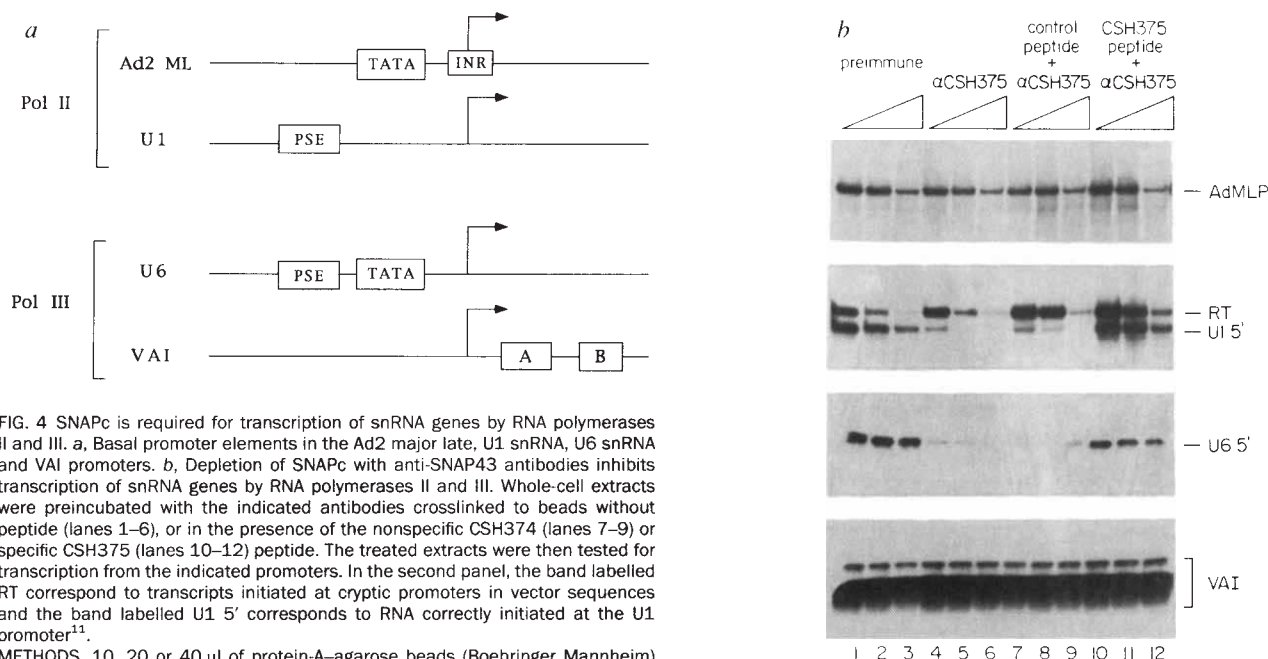


FIG. 4 SNAPc is required for transcription of snRNA genes by RNA polymerases II and III. **a**, Basal promoter elements in the Ad2 major late, U1 snRNA, U6 snRNA and VAI promoters. **b**, Depletion of SNAPc with anti-SNAP43 antibodies inhibits transcription of snRNA genes by RNA polymerases II and III. Whole-cell extracts were preincubated with the indicated antibodies crosslinked to beads without peptide (lanes 1–6), or in the presence of the nonspecific CSH374 (lanes 7–9) or specific CSH375 (lanes 10–12) peptide. The treated extracts were then tested for transcription from the indicated promoters. In the second panel, the band labelled RT correspond to transcripts initiated at cryptic promoters in vector sequences and the band labelled U1 5' corresponds to RNA correctly initiated at the U1 promoter¹¹.

METHODS. 10, 20 or 40 μ l of protein-A-agarose beads (Boehringer Mannheim) crosslinked to antibodies from preimmune serum (lanes 1–3) or anti-CSH375 antibodies (lanes 4–12) were incubated with 40 μ l HeLa wholecell extract¹¹ (25 mg protein per ml) for 1 h at room temperature. In lanes 7–9 and 10–12, 40 μ g nonspecific peptide CSH374 and specific peptide CSH375 (Fig. 2), respectively, were added to the extract 10 min before addition of the antibody

beads. 4, 15, 4 and 4 μ l of the depleted extracts were then tested in parallel for transcription from the Ad2 major late promoter [p119MLP(C2A)]⁷, the U1 promoter (pU1*G)¹¹, the U6 promoter (pU6*/Hae/RA.2)¹¹, and the VAI promoter (pSBM13*VAI)⁷, respectively, as described^{7,11}.

Because highly purified SNAPc is unstable and TBP in particular seems to dissociate from the complex, we incubated crude nuclear extract with beads crosslinked to anti-CSH375 antibodies or to preimmune antibodies, and eluted the bound proteins with the specific CSH375 peptide. As shown in Fig. 3a, the SNAP complex was eluted from the anti-CSH375 beads (lanes 2, 3), but not from the preimmune beads (lanes 4, 5). The eluted material was then probed in an immunoblot with anti-TBP antibodies, stripped and reprobed with the anti-CSH375 antibodies (Fig. 3b). Like SNAP43, TBP was present in the material eluted from the anti-CSH375 beads, but not in the material eluted from preimmune beads (compare lanes 3 and 4). Thus, TBP was specifically co-immunoprecipitated with SNAP43, indicating that TBP is not recruited into the SNAP complex upon binding of SNAPc to the PSE, but rather is a genuine component of free SNAPc. The interaction between TBP and SNAP43 may be direct, as SNAP43 is retained efficiently on a glutathione-agarose column coated with a GST-TBP fusion protein, but not on control columns coated with an irrelevant fusion protein or GST alone (Fig. 3c).

We used the ability of the anti-CSH375 antibodies to recognize SNAPc in the absence of its DNA-binding site to deplete extracts of the complex and so determine the SNAPc requirement for transcription from four different promoters, whose basal elements are shown in Fig. 4a. In the RNA polymerase II adenovirus 2 (Ad2) major late and human U1 snRNA promoter, basal transcription is directed by a TATA box combined with an initiator and by a PSE, respectively. In the RNA polymerase III U6 snRNA and VAI promoters, basal transcription is directed by a PSE combined with a TATA box and by A and B boxes inside the gene, respectively. As shown in Fig. 4b, depletion with anti-CSH375 beads resulted in much stronger inhibition of U1 and U6 transcription than depletion with beads coated with preimmune antibodies (compare lanes 4–6 with 1–3). This inhibitory effect could be blocked by addition of the specific peptide antigen CSH375, but not by addition of an irrelevant peptide (peptide CSH374) (compare lanes 10–12 with 7–9). In contrast, the depletions had no specific effect on transcription from the Ad2 major late and VAI promoters. Thus, the TBP-containing complex SNAPc is required specifically for

both RNA polymerase II and III transcription of snRNA genes. So far, we have not observed a SNAPc requirement for transcription from other classes of promoters.

SNAPc has unusual DNA-binding characteristics in that it binds specifically to a non-TATA-box DNA-binding site, the PSE. Although we do not know whether TBP contacts DNA in the complex, it is clear that TBP is not essential for DNA binding, because depletion of TBP from the SNAP complex with anti-TBP antibodies results in the appearance of smaller complexes that can still bind specifically to the PSE¹¹. SNAP43 on its own has no detectable DNA-binding activity in an electrophoretic mobility-shift assay, consistent with the absence of DNA-binding motifs in the predicted amino-acid sequence and with the observation that SNAP43, unlike other members of the SNAP complex, cannot be crosslinked to DNA (unpublished results). Thus, other members of the SNAP complex probably bind specifically to the PSE and, as they are associated with TBP, in effect recruit TBP to a non-TATA-box sequence. □

Received 14 October 1994; accepted 9 February 1995.

- Hernandez, N. *Genes Dev.* **7**, 1291–1308 (1993).
- Comai, L., Tanese, N. & Tjian, R. *Cell* **68**, 965–976 (1992).
- Goodrich, J. A. & Tjian, R. *Curr. Opin. Cell Biol.* **6**, 403–409 (1994).
- Tjian, R. & Maniatis, T. *Cell* **77**, 5–8 (1994).
- Huet, J. & Sentenac, A. *Nucleic Acids Res.* **20**, 6451–6454 (1992).
- Kassavetis, G. A. et al. *Cell* **71**, 1055–1064 (1992).
- Lobo, S. M., Tanaka, M., Sullivan, M. L. & Hernandez, N. *Cell* 1029–1040 (1992).
- Simmen, K. A., Bernues, J., Lewis, J. D. & Mattaj, J. W. *Nucleic Acids Res.* **20**, 5889–5898 (1992).
- Taggart, A. K. P., Fisher, T. S. & Pugh, B. F. *Cell* **71**, 1015–1028 (1992).
- White, R. J. & Jackson, S. P. *Cell* 1041–1053 (1992).
- Sadowski, C. L., Henry, R. W., Lobo, S. M. & Hernandez, N. *Genes Dev.* **7**, 1535–1548 (1993).
- Hernandez, N. in *Transcriptional Regulation* 281–313 (Cold Spring Harbor Monograph Series, Cold Spring Harbor, New York, 1992).
- Lobo, S. & Hernandez, N. in *Transcription, Mechanisms and Regulation* (eds Conaway, R. C. & Conaway, J. W.) 127–159 (Raven, New York, 1994).
- Xiong, Y. et al. *Nature* **366**, 701–704 (1993).
- Skowronski, J. et al. *Molec. cell. Biol.* **8**, 1385–1397 (1988).
- Niman, H. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4949–4953 (1983).

ACKNOWLEDGEMENTS. We thank M. Sullivan and B. Ma for technical help; W. Herr, D. Morrison, J. Skowronski, M. Tanaka and A. Wilson for reagents; D. Marshak and G. Binns for peptide synthesis; S. Bell, W. Chang, K. Fien and D. Marshak for discussion and for advice; W. Herr, P. Kaufman and B. Stillman for comments on the manuscript; R. Frenguelli for secretarial help; and J. Duffy, M. Ockler and P. Renna for artwork and photography. This work was supported by a grant from the NIH to N.H.

Basal promoter elements as a selective determinant of transcriptional activator function

Gokul Das*, Craig S. Hinkley & Winship Herr

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

IN eukaryotes, activation of transcription involves an interplay between activators bound to *cis*-regulatory elements and factors bound to basal elements near the start site of transcription. The basal elements, for example the TATA box or proximal sequence element (PSE) of small nuclear RNA (snRNA) promoters, nucleate the assembly of basal transcription complexes, components of which interact with activators^{1,2}. Although one basal transcription complex can interact with many activators, it is unclear whether different basal transcription complexes can direct different responses to particular activators. We show here that changing the arrangement of basal elements can alter the response to transcriptional activation domains. Indeed, in the human U6 snRNA promoter, point mutation of either a TATA box or PSE results in diametrically opposed responses to VP16- and Sp1-derived activation domains. These basal elements can even discriminate small changes in an activation domain. Thus the arrangement of basal promoter elements provides a mechanism for differential regulation of transcription.

To study the response of basal elements to various activation domains, nine different activation domains were fused to the yeast GAL4 DNA-binding domain (residues 1 to 94), as illustrated in Fig. 1, and their activity assayed in different promoter contexts. Figure 2 shows the *in vivo* response of two RNA polymerase II promoters, the c-Fos messenger RNA and U2 snRNA promoters, to the nine different activators. As illustrated in Fig. 2b, the c-Fos promoter contains a TATA box, which typically binds the TATA-binding protein (TBP)-containing TFIID complex³, whereas the U2 promoter contains a PSE, which binds a different TBP-containing complex called SNAP⁴. The U2 reporter construct generates a correctly initiated U2 snRNA transcript (labelled U2 in Fig. 2a) and a 'read-through' transcript (labelled U2 RT), which is polyadenylated and probably results from a cryptic mRNA-type 'U2 RT' promoter in the reporter construct⁵. This reporter construct can therefore measure simultaneously the responses of snRNA- and mRNA-type promoters to different activation domains.

As expected^{6,8}, the acidic GAL4-VP16 protein was a potent activator of the c-Fos mRNA promoter and U2 RT promoter (Fig. 2a, lanes 1–3). The glutamine-rich GAL4-Sp1^Q protein, however, was 50- to 100-fold less active on these two promoters, although still 3-fold more active than the GAL4 DNA-binding domain alone (compare lanes 2 and 7). In striking contrast, like the weaker GAL4 activation domain⁸, the otherwise very potent VP16 activation domain displayed little if any activation of the U2 snRNA promoter, whereas the Sp1^Q activation domain was the strongest U2 snRNA promoter activator (lanes 3–7). The inability of the VP16-related activator to stimulate the U2 snRNA promoter was not due to occlusion by the very active U2 RT promoter, because the GAL4-Sp1^Q activator was able to activate the U2 snRNA promoter in the presence of the GAL4-VP16 activator (data not shown).

Previous studies⁹ have shown that activation domains in Oct-1 and Oct-2 also display promoter selectivity; those in Oct-1 are

better activators of the U2 snRNA promoter, whereas those in Oct-2 are better activators of the β -globin mRNA promoter. We tested whether activation domains composed of four tandem copies of 19- or 18-amino-acid segments derived from glutamine-rich activation domains in Oct-1 and Oct-2 (called Q¹⁹ and Q¹⁸, respectively; Fig. 1b) possess the promoter selectivity of their parents (Fig. 2a; lanes 8–11). Indeed, the Oct-2-derived 4 × Q¹⁸ activation domain is a better activator of the c-Fos and U2 RT mRNA-type promoters (see also ref. 10), whereas the Oct-1-derived 4 × Q¹⁹ activation domain is a better activator of the U2 snRNA promoter (compare lanes 8 and 9). Thus even these very short segments, when reiterated, can reproduce the promoter selectivities of the parental Oct-1 and Oct-2 proteins. Curiously, although two mutant Q¹⁸ segments were debilitated for activation of the c-Fos and U2 RT mRNA-type promoters as expected¹⁰, the L→A (lacking leucines), but not the Q→A (lacking glutamines), mutant Q¹⁸ segment was active on the U2 snRNA promoter (lanes 10 and 11).

Although there are many other differences between these promoters, it is likely that TATA box and PSE basal promoter elements direct the differential responses of the c-Fos and U2 promoters to activation domains. The human U6 snRNA pro-

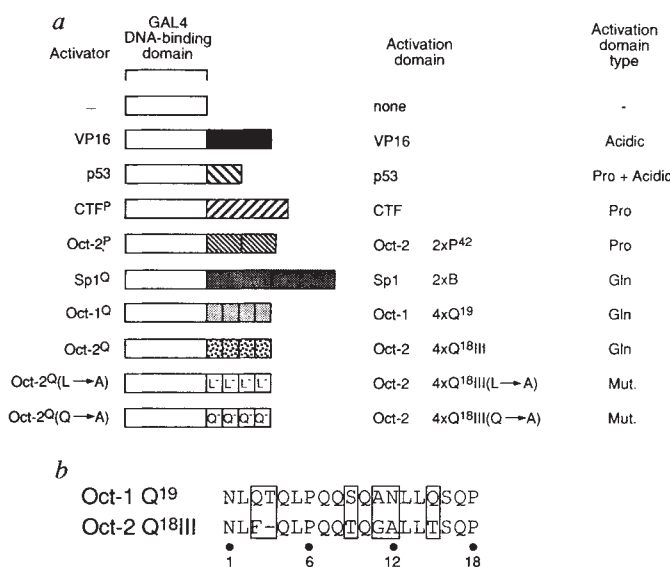


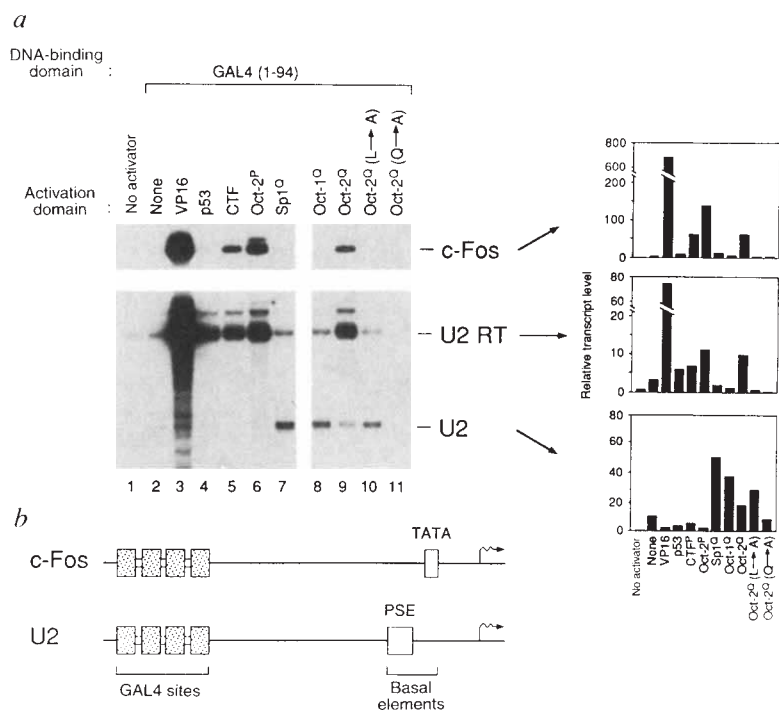
FIG. 1 Transcriptional activator structure. **a**, Activator proteins. The GAL4 DNA-binding domain (residues 1 to 94) was fused to one viral and eight human transcription factor-derived activating regions: VP16, residues 413 to 490 of the herpes simplex virus protein VP16 (ref. 21); p53, residues 1 to 42 of the anti-oncoprotein p53 (ref. 22); CTF, carboxy-terminal residues 399 to 499 of the CCAAT-box transcription factor CTF²³; Oct-2(2 × P⁴²), carboxy-terminal Oct-2 residues 438 to 479 (ref. 24); Sp1(2 × B), two tandem copies of residues 263 to 391 of the Sp1 B domain²⁵; Oct-1 (4 × Q¹⁹), four tandem copies of Oct-1 residues 223 to 241 (ref. 10); Oct-2(4 × Q¹⁸), four tandem copies of Oct-2 residues 143 to 160 (ref. 10); Oct-2(4 × Q¹⁸L→A) and Oct-2(4 × Q¹⁸Q→A), matched mutant versions of Oct-2(4 × Q¹⁸) where all leucine (L→A) or glutamine (Q→A) residues in all four copies have been mutated to alanine¹⁰. The Sp1(2 × B) activation domain is not shown to scale. The activation domain type is indicated. (Mut., mutant). **b**, Sequence comparison of Oct-1 Q¹⁹ and Oct-2 Q¹⁸III activation domain segments. Sequence differences are boxed.

METHODS. All activator expression plasmids are derivatives of the CMV promoter-driven GAL4 DNA-binding domain expression construct pCG-GAL4 (1 to 94) (ref. 24) and contain the activation domain sequences inserted between the unique 5' XbaI and 3' BamHI sites of pCG-GAL4(1 to 94). The Sp1(2 × B) and CTF (ref. 26), Oct-2(2 × P⁴²) (ref. 24) and Oct-2(4 × Q¹⁸) and Oct-2(4 × Q¹⁸L→A) (ref. 10) pCG-GAL4 constructs have been described previously. The VP16, p53 and 4 × Q¹⁹ pCG-GAL4 constructs were prepared for this study. The pCG-GAL4-Oct-2(4 × Q¹⁸Q→A) expression construct was a gift from M. Tanaka (M. Tanaka and W.H., unpublished data).

* Present address: Cancer Therapy & Research Center and Department of Cellular and Structural Biology, University of Texas Health Science Center, 8122 Datapoint Drive, San Antonio, Texas 78229, USA

FIG. 2 Transcriptional activation specificity of nine activation domains on the c-Fos mRNA and U2 snRNA promoters. **a**, RNase protection assay of reporter transcripts from the c-Fos and U2 promoter constructs after transient expression in HeLa cells. Transcript levels are shown relative to the sample without activator; α -globin reference fragments are not shown. U2 RT, U2 read-through transcript. **b**, Illustration of the c-Fos and U2 snRNA promoters. The two promoters contain four tandem GAL4-binding sites (G17M²⁷; stippled boxes) inserted at position -197 (c-Fos promoter) or -198 (U2 promoter) upstream of the transcriptional start sites (wavy arrows).

METHODS. HeLa cells (1×10^6) were transfected by calcium phosphate coprecipitation with either 2 μ g (c-Fos) or 4 μ g of target promoter construct, either 320 ng (pCG-GAL4(1 to 94)) or 640 ng of transactivator expression plasmid, and 66 ng of reference plasmid p $\alpha 4 \times (A + C)$ ²⁸. Cytoplasmic RNA and protein were prepared²⁹ and RNA expression analysed by RNase protection⁹ as described for c-Fos²⁴, U6 and U2 (ref. 12), and internal control α -globin transcripts. Quantitation was with a Fuji BAS1000 Bioimaging Analyzer System. The quality and quantity of the expressed activator proteins were monitored by immunoblotting with GAL4 antisera and electrophoretic mobility retardation analysis (data not shown). All of the activators were expressed as single, apparently full-length species, but two activators, the GAL4-p53 and especially the GAL4-VP16 activator, were expressed at lower levels, but this lower expression level did not prevent transcriptional activation. The c-Fos reporter pc-fos/-197/4 $\times$ G17M was created by insertion of the GAL4 binding-site-containing *Sall*-*XbaI* fragment from pc-fos/-56/4 $\times$ G17M²⁴ between the *Sall* and *XbaI* sites in pc-fos/-197/+109 (ref. 30). pU2/-198/4 $\times$ G17M was constructed by insertion of a 4 $\times$ G17M GAL4 binding-site-containing



EcoRI-*HincII* fragment from a pUC119 clone²⁴ between the *EcoRI* and *SmaI* sites in U2/-247 (ref. 12).

FIG. 3 Basal transcriptional elements in the human U6 promoter regulate the response to transcriptional activation domains. **a**, RNase protection assay of reporter transcripts from the U6 wild-type, TATA⁻ and PSE⁻ promoter construct after transient expression in HeLa cells. **b**, Illustration of the three U6-related promoters, and their transcriptional characteristics and activation profiles. The four GAL4 binding sites are inserted at position -198 from the start site in the wild-type U6 promoter. The U6 TATA⁻ and PSE⁻ promoters only differ by the presence of the TATA box (Is7) and PSE (Is2) point mutations¹². The position of the start site in the mutant U6 promoters varies slightly: in the TATA⁻ promoter it is approximately three to four base pairs upstream¹² and in the PSE⁻ promoter four to five base pairs downstream of the wild-type start site (data not shown). The type of transcripts from each promoter was ascertained by characterizing their 3' ends and by sensitivity to α -amanitin as previously described¹² (data not shown). Quantitation of the transcript levels in **a** is shown and was determined as described in Fig. 2 legend.

METHODS. The transient expression, RNase protection and activator expression assays were performed as described in the Fig. 2 legend. To construct the GAL4 binding-site-containing wild-type U6 promoter plasmid pU6/-198/4 $\times$ G17M, a unique *EcoRV* site was created at position -198 of pU6/Hae/RA.2 (ref. 12), generating pU6/Hae/RA.2/*EcoRV*, and a *SacI*-*HincII* restriction fragment containing four G17M GAL4-binding sites from a pUC119 clone²⁴ was inserted between the *SacI* and *EcoRV* sites of pU6/Hae/RA.2/*EcoRV*. pU6/PSE⁻/im198/4 $\times$ G17M and pU6/TATA⁻/im198/4 $\times$ G17M were created by substitution of the wild-type U6 promoter sequences for the mutant U6 Is2 (PSE⁻) and Is7 (TATA⁻) promoter sequences described previously¹².

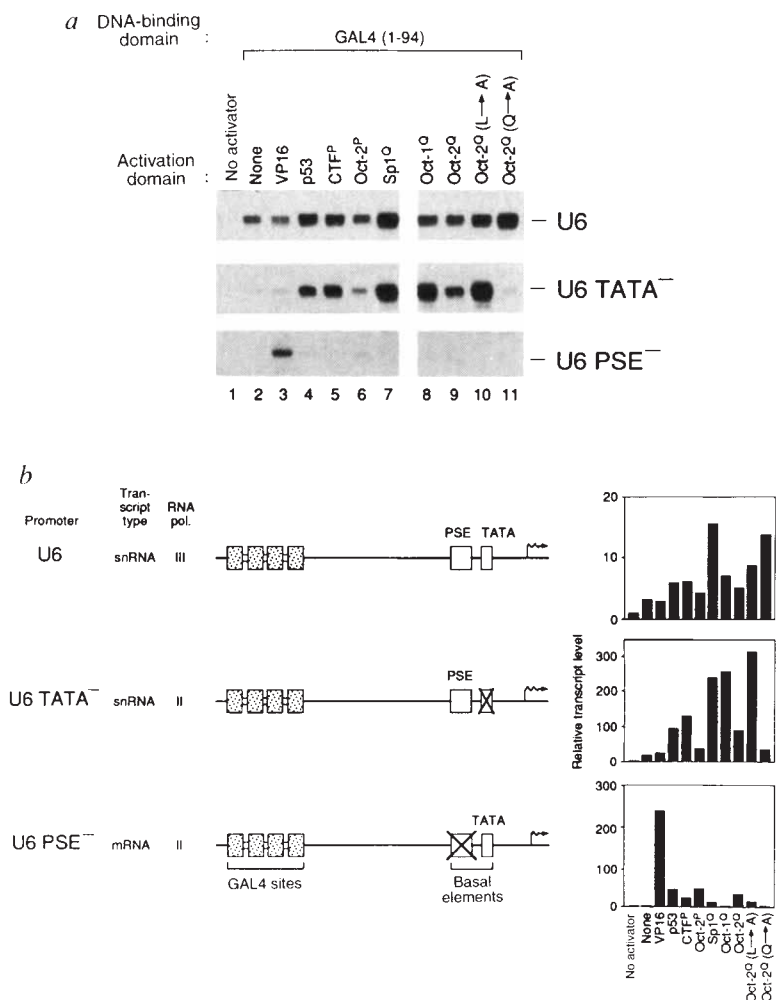
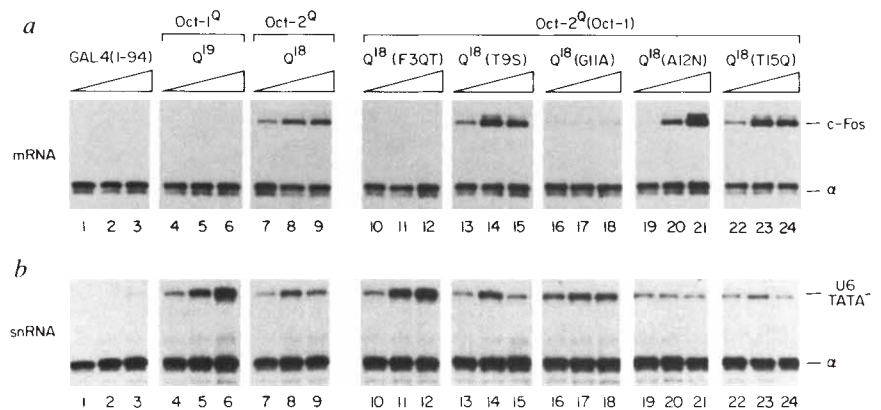


FIG. 4 Small changes within an activation domain can alter its promoter-selective properties. RNase protection assay of *a*, c-Fos and *b*, U6 TATA reporter transcripts after transient expression in HeLa cells. The identity of the Oct-1^Q and Oct-2^Q activator constructs, and Oct-2^Q (Oct-1) derivatives carrying the Oct-1-derived replacements (refer to Fig. 1b) is indicated above the lanes. For example, the Q¹⁸(F3QT) segment carries the Oct-1-derived QT sequence in place of the single Oct-2 phenylalanine (F) residue. Amino-acid substitutions are present in each of the four tandemly reiterated Q¹⁸ segments of these activators. Each panel of three lanes shows a titration of increasing amounts of effector plasmid (0.375 µg, 0.75 µg or 1.5 µg). The identity of the protected fragments from correctly initiated c-Fos, U6 TATA⁻ and α -globin internal control transcripts is indicated to the right.

METHODS. HeLa cells ($3-4 \times 10^6$) were transfected by electroporation with 4 µg promoter construct, 80 ng reference plasmid pA4 × (A + C), and 0.375 µg, 0.75 µg or 1.5 µg effector expression plasmid. RNA was analysed by RNase protection as described in Fig. 2 legend, and protein



was analysed by electrophoretic mobility retardation and immunoblot analyses (data not shown). In each case, full-length proteins were expressed at levels within twofold of each other at the highest amount of activator plasmid transfected (1.5 µg).

moter is ideal to test this hypothesis, because it contains both a TATA box and a PSE (see Fig. 3b), which in combination direct transcription by RNA polymerase III. Point mutations, however, in either the U6 TATA box (U6 TATA⁻) or PSE (U6 PSE⁻) result in RNA polymerase II transcripts of either the U2 snRNA-type^{11,12} or mRNA-type (see Fig. 3 legend), respectively. Thus, by comparing the ability of the wild-type, TATA⁻, and PSE⁻ U6 promoters to respond to activators, the role of basal elements in the control of activation domain activity can be assessed.

Figure 3 shows such a comparison. Curiously, although the wild-type U6 promoter contains a TATA box, it fails to respond effectively to the otherwise very potent VP16 activation domain; instead it responds effectively to the Sp1^Q activation domain (Fig. 3a, compare lanes 2, 3 and 7). Indeed, the wild-type RNA polymerase III U6 promoter responds to the battery of activation domains in a pattern similar (albeit not identical) to that of the RNA polymerase II U2 snRNA promoter (compare Figs 2 and 3), indicating that different RNA polymerases can have similar activation domain response patterns.

Except for the p53- and CTF-related activators, the U6 TATA⁻ promoter and the U2 promoter respond to activators in a similar way (Fig. 2a and 3a), consistent with their shared RNA polymerase II snRNA-type transcription properties¹². These promoters all possess an snRNA-specific PSE, suggesting that this basal element has a profound influence on the response of a promoter to activation domains. Indeed, when the PSE is mutated and the TATA box is retained, as in the U6 PSE⁻ promoter, the activation domain response pattern is very different (compare panels in Fig. 3a). Now, although weaker, the response pattern is similar to that of the c-Fos mRNA promoter (cf. Fig. 2a), where the VP16-derived activator is more active than the Sp1-derived activator. Thus, by changing the arrangement of basal promoter elements, a single RNA polymerase can exhibit very different activation domain response patterns.

Together, the response of the various mRNA and snRNA promoters to the mutant Oct-2^Q (Q→A) activation domain (Figs 2 and 3) reveals a startling result. Although the Oct-2^Q (Q→A) activation domain displays little, if any, activity on the RNA polymerase II c-Fos, U2 and U6 PSE⁻ and TATA⁻ promoters, it is very active on the wild-type U6 promoter (Fig. 3a, lane 11). Thus a combination of two different basal elements specifies both RNA polymerase III transcription and a unique activation domain response pattern.

Figure 4 examines the sensitivity of basal elements to differences in activation domain structure, by determining the influence of the five differences between the Q¹⁹ and Q¹⁸ segments

(see Fig. 1b) on the different promoter selective properties. Residues in the Q¹⁸ segment were individually replaced by their Oct-1 counterpart(s) and the replacements assayed on the c-Fos mRNA-type (Fig. 4a) and U6 TATA⁻ snRNA-type (Fig. 4b) promoters. Three of the amino-acid exchanges, T9S, A12N and T15Q, have little effect on the promoter selectivity of the reiterated Q¹⁸ activation domain (compare lanes 3–15 and 19–24 with lanes 4–6). The G11A exchange has an intermediate effect (lanes 16–18), but the most dramatic effect, fully reversing the activity of the Q¹⁸ segment, is exhibited by the F3QT exchange (compare lanes 10–12 with lanes 4–9). The F3QT exchange also reverses the activity of the Q¹⁸ segment on the natural U2 and U6 promoters (data not shown).

These results show that replacement of a single residue, a phenylalanine, within a reiterated segment by two other residues can greatly influence the promoter selectivity of an activation domain. A phenylalanine residue in the VP16 activation domain is also critical for activation of an mRNA promoter¹³; these two activation domains may use related secondary structures or hydrophobic surfaces to interact with targets in an mRNA promoter transcription complex that are not necessary for interaction with an snRNA transcription complex.

The experiments described here demonstrate that the arrangement of basal promoter elements, even in the context of identical activator binding sites, can profoundly influence the response of a promoter to different activators. This finding provides a mechanism by which different promoters can display very different responses to a common set of activators, in particular activators with the same DNA-binding specificity. These different activities may result from different interactions, either direct or indirect^{2,14,20}, between activation domains and the basal transcriptional machinery assembled on these promoters. The robust differential response observed *in vivo* indicates that control of activator specificity mediated by basal promoter elements is one of the mechanisms by which genes can selectively respond to the multitude of activators expressed in a cell. □

Received 21 December 1994; accepted 17 February 1995.

- Hernandez, N. *Genes Dev.* **7**, 1291–1308 (1993).
- Tjian, R. & Maniatis, T. *Cell* **77**, 5–8 (1994).
- Nakajima, N., Horikoshi, M. & Roeder, R. G. *Molec. cell. Biol.* **8**, 4028–4040 (1988).
- Sadowski, C. L., Henry, R. W., Lobo, S. M. & Hernandez, N. *Genes Dev.* **7**, 1535–1548 (1993).
- Hernandez, N. & Lucito, R. *EMBO J.* **7**, 3125–3134 (1988).
- Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. *Nature* **335**, 563–564 (1988).
- Cousens, D. J., Greaves, R., Goding, C. R. & O'Hare, P. *EMBO J.* **8**, 2337–2342 (1989).
- Tanaka, M., Grossniklaus, U., Herr, W. & Hernandez, N. *Genes Dev.* **2**, 1764–1778 (1988).
- Tanaka, M., Lai, J.-S. & Herr, W. *Cell* **68**, 755–767 (1992).
- Tanaka, M. & Herr, W. *Molec. cell. Biol.* **14**, 6056–6067 (1994).

11. Mattaj, J. W., Dathan, N. A., Parry, H. D., Carbon, P. & Krol, A. *Cell* **55**, 435–442 (1988).
12. Lobo, S. M. & Hernandez, N. *Cell* **58**, 55–67 (1989).
13. Cress, W. D. & Triezenberg, S. J. *Science* **251**, 87–90 (1991).
14. Stringer, K. F., Ingles, C. J. & Greenblatt, J. *Nature* **345**, 783–786 (1990).
15. Choy, B. & Green, M. R. *Nature* **366**, 531–536 (1993).
16. Brou, C. *et al.* *EMBO J.* **12**, 489–499 (1993).
17. Koleske, A. J. & Young, R. A. *Nature* **368**, 466–469 (1994).
18. Kim, Y.-J., Björklund, S., Li, Y., Sayre, M. H. & Kornberg, R. D. *Cell* **77**, 599–608 (1994).
19. Kim, T. K. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **91**, 4170–4174 (1994).
20. Chen, J.-L., Attardi, L. D., Verrijzer, C. P., Yokomori, K. & Tjian, R. *Cell* **79**, 93–105 (1994).
21. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. *Genes Dev.* **2**, 718–729 (1988).
22. Unger, T., Nau, M. M., Segal, S. & Minna, J. D. *EMBO J.* **11**, 1383–1390 (1992).
23. Mermod, N., O'Neill, E. A., Kelly, T. J. & Tjian, R. *Cell* **58**, 741–753 (1989).
24. Tanaka, M., Clouston, W. M. & Herr, W. *Molec. cell. Biol.* **14**, 6046–6055 (1994).
25. Courey, A. J. & Tjian, R. *Cell* **55**, 887–898 (1988).
26. Tansey, W. P., Ruppert, S., Tjian, R. & Herr, W. *Genes Dev.* **8**, 2756–2769 (1994).
27. Webster, N., Jin, J. R., Green, S., Hollis, M. & Chambon, P. *Cell* **52**, 169–178 (1988).
28. Tanaka, M. & Herr, W. *Cell* **60**, 375–386 (1990).
29. Das, G. & Herr, W. *J. biol. Chem.* **268**, 25026–25032 (1993).
30. Gilman, M. Z., Wilson, R. N. & Weinberg, R. A. *Molec. cell. Biol.* **6**, 4305–4316 (1986).

ACKNOWLEDGEMENTS. We thank M. Tanaka and W. Tansey for effector expression constructs; M. Gilman and M. Tanaka (c-fos), S. Lobo and N. Hernandez (U6 and U2), C. Prives (p53), and S. Triezenberg (VP16) for clones; and I. Sadowski for GAL4 antisera. We also thank S. Bell, M. Cleary, N. Hernandez, R. Singh, M. Tanaka and W. Tansey for comments on the manuscript; and J. Duffy, M. Ockler and P. Renna for artwork and photography. G.D. was a fellow of the Leukemia Society of America, and C.S.H. was supported by an American Cancer Society fellowship. These studies were supported by a grant from the National Cancer Institute.

RNA polymerase II C-terminal domain required for enhancer-driven transcription

Hans-Peter Gerber, Michael Hagmann,
Katja Seipel, Oleg Georgiev, Marilyn A. L. West*,
Ying Litigting†, Walter Schaffner
& Jeffry L. Corden†

Institut für Molekularbiologie II der Universität Zürich,
Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

* Section of Plant Biology, Division of Biological Science,
University of California, Davis, California 95616, USA

† Howard Hughes Medical Institute and Department of Molecular
Biology and Genetics, Johns Hopkins University School of Medicine,
Baltimore, Maryland 21205, USA

THE RNA polymerase II carboxy-terminal domain (CTD) consists of tandem repeats of the sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser^{1–3}. The CTD may participate in activated transcription through interaction with a high-molecular-weight mediator complex^{4,6}. Such a role would be consistent with observations that some genes are preferentially sensitive to CTD mutations^{7,8}. Here we investigate the function of the mouse RNA polymerase CTD in enhancer-driven transcription. Transcription by α -amanitin-resistant CTD-deletion mutants was tested by transient transfection of tissue culture cells in the presence of α -amanitin in order to inhibit endogenous RNA polymerase II. Removal of most of the CTD abolishes transcriptional activation by all enhancers tested, whereas transcription from promoters driven by Sp1, a factor that typically activates housekeeping genes from positions proximal to the initiation sites, is not affected. These findings show that the CTD is essential in mediating 'enhancer'-type activation of mammalian transcription.

To study functional interactions between enhancer/promoter elements and the CTD of mammalian RNA polymerase II (pol II), we transiently transfected tissue culture cells with α -amanitin-resistant⁹ mutant genes encoding the pol II largest subunit and reporter plasmids containing different enhancer and/or promoter proximal elements. Transcription by mutant pol II was assessed by inhibiting the endogenous enzyme with α -amanitin. This strategy has several advantages over other previously used approaches. Yeast CTD truncation mutants reduce GAL4-activated transcription to 50% of the wild-type value⁸; the minor extent of reduction is probably due to the retention of enough

CTD repeats for viability. Our transient expression assay allows us to study the transcription capacity of otherwise lethal CTD truncations *in vivo*, which is important because *in vitro* transcription activation by enhancers is relatively weak¹⁰ and the entire CTD may be dispensable altogether *in vitro*^{11,12}.

Vectors used for transfection expressed CTD mutants containing all 52 heptad repeats (wild type, WT), 31 repeats (Δ 31) or only 5 repeats (Δ 5) (Fig. 1a). The WT and Δ 31 constructs are able to support long-term growth, whereas the Δ 5 mutant does not¹³. Western blot analysis of nuclear extract from transiently transfected monkey COS-1 cells shows that the three α -amanitin-resistant RNA pol II large subunits were present in the nucleus (Fig. 1b). Whole-cell extracts from transfected cells incubated in the presence of α -amanitin for 40 h contained as much or more of the mutant large subunit (compare with WT; Fig. 1c), indicating steady-state levels of the mutant polymerases were comparable during the time of the experiment. To determine whether the overall condition of cells expressing the three CTD forms was differentially affected during α -amanitin incubation, we cotransfected the pol III-transcribed VAI RNA gene and analysed transcription by S1 nuclease mapping¹⁴. Although there was a general decline in pol III transcription, it was the same for cells containing the three different pol II forms (Fig. 1d, lanes 3 to 5). The combined results of these control experiments indicate that treatment of transfected cells with α -amanitin does not differentially affect cells expressing mutant RNA polymerase subunits during the course of the experiments.

To assess the effect of CTD truncation on expression driven by promoter proximal elements, we cotransfected the three mutant forms of the largest subunit of RNA pol II with reporter constructs driven by tandem Sp1 sites from a herpes simplex virus promoter. In addition, we tested Sp1 binding sites present in the 21-base-pair (bp) repeats of the SV40 early promoter region (for constructions, see Fig. 2a). As demonstrated in Fig. 2b, cotransfection of any of the mutant forms of α -amanitin-resistant largest subunits confers α -amanitin-resistant transcription to the different Sp1 reporter constructs tested (lanes 2–4), indicating that all three CTD forms are assembled into functional RNA pol II and can transcribe the two Sp1-driven reporter genes with equal efficiency.

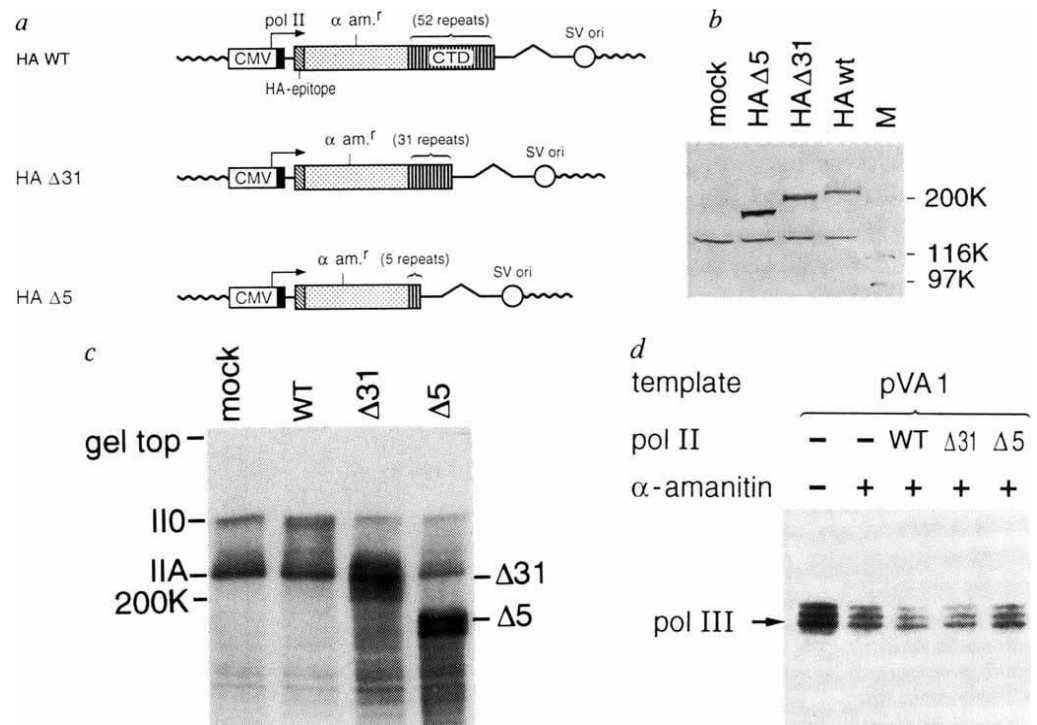
To test enhancers, we cotransfected a target promoter template that contains the SV40 enhancer together with each of the three pol II largest-subunit forms (for constructions, see Fig. 3a). We observed a dramatic reduction in SV40 enhancer-driven transcription with the most severe CTD truncation and a moderate (twofold) reduction in Δ 31. Transcription driven by the Sp1 promoter was not affected (compare lanes 6 and 8 in Fig. 3b). Findings were similar in Δ 5-transfected cells for the metal-responsive enhancer (MRE), the mouse mammary tumour virus (MMTV) enhancer, the immunoglobulin heavy-chain enhancer (IgH), the GAL4 enhancer driven by a GAL4-VP16 fusion protein (Fig. 3c), and the cytomegalovirus (CMV) enhancer (not shown). The reduction in SV40 enhancer-driven transcription by Δ 5 is the same irrespective of whether the enhancer is located 5' or 3' of the promoter and in the presence or absence of Sp1 sites (compare Fig. 3b and d). We conclude that the efficient enhancer-boosted transcription is more sensitive to CTD truncation than transcription driven by a promoter proximal element.

The inability of pol II Δ 5 to respond to enhancers is seen even at short α -amanitin incubation times (Fig. 3d: compare lanes 3, 6 and 9 with lanes 2, 5 and 8, respectively). This result suggests that the inability to respond to an enhancer is a direct result of CTD truncation.

The results presented here are consistent with a function for the CTD in mediating enhancer activity and support earlier evidence from yeast implicating the CTD in GAL4 function^{7,8}. Removal of almost all of the mammalian CTD has a more profound effect on enhancer function, with reductions of more than tenfold for every enhancer tested. Surprisingly, transcription driven by promoter-proximal Sp1 binding sites is unaffected.

FIG. 1 Expression of RPII215 CTD forms. **a**, Expression vectors used in transfection experiments with HeLa or COS-1 cells. Genomic RPII215 DNA was cloned into the eukaryotic expression vector pSTC, driven by a strong cytomegalovirus (CMV) promoter²⁰. The pol II $\Delta 31$ and $\Delta 5$ constructs included 31 or 5 heptad repeats, respectively, and correspond to the original clones A21, Del 23–36+39–47 and Del2–50', respectively¹³. Initiation of translation occurs at the AUG codon of the thymidine kinase leader sequence followed by five additional amino acids before the RPII215 sequence. Constructs HAWT, HA $\Delta 31$ and HA $\Delta 5$ contain nine additional amino acids forming the haemagglutinin epitope (as given above) at the N terminus of the protein. **b**, The RPII215 constructs are present in the nucleus of transfected COS-1 cells. 10 μ g nuclear extract¹⁴ of COS-1 cells lipofected with 10 μ g of the three RPII215-CTD constructs indicated were loaded on a 5% SDS-

polyacrylamide gel. The nitrocellulose filter was incubated with a monoclonal antibody (12CA5, BAbCO) directed against the 9-amino-acid haemagglutinin epitope (Tyr-Pro-Tyr-Asp-Val-Pro-Asp-Tyr-Ala) fused to the N terminus of RPII215 CTD constructs (HA $\Delta 5$, HA $\Delta 31$ and HAWT). **c**, RPII215 constructs are expressed in the presence of α -amanitin. Whole-cell extracts were prepared from COS-1 cells by boiling transfected cells in SDS-sample buffer. Proteins were separated by 5% SDS-PAGE and electrophoretically transferred to nitrocellulose. The filter was incubated with a monoclonal antibody directed against the RPII215 largest subunit (ARNA-3; Progen) and detected by ECL (Amersham). Because the cells are rapidly lysed both the phosphorylated and unphosphorylated forms of the largest subunit are observed.



Positions of the wild-type phosphorylated (IIO) and unphosphorylated (IIA) wild-type large subunits are shown to the left; the positions of the deletion subunits are shown to the right. We observe two bands for $\Delta 5$, indicating that even though only five heptapeptide repeats remain, some CTD phosphorylation takes place. **d**, Pol III transcription is not differentially affected in the presence of any of the RPII215 CTD forms during the standard amanitin-treatment period.

METHODS. COS-1 cells were lipofected with 3 μ g pVA1 and 10 μ g of the indicated RPII215-CTD constructs for 5 h and grown for an additional 10 h before α -amanitin (2.5 μ g ml⁻¹) was added for an additional 40 h. Total RNA (10 μ g) was analysed by S1 mapping¹⁴. Western blotting conditions have been described²¹.

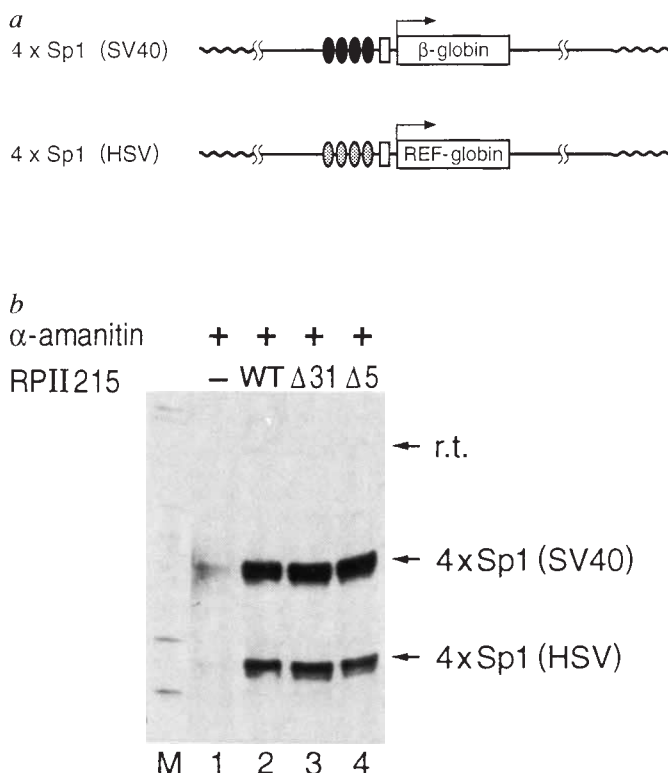


FIG. 2 All RPII215-CTD forms yield α -amanitin-resistant transcription from two different Sp1 promoters. **a**, All pol II reporter constructs described here and in Fig. 3 included the same β -globin minimal promoter consisting of the TATA box (CATATAA) and the β -globin initiation site²². Filled circles, Sp1 sites derived from the SV40 enhancer; stippled circles, Sp1 sites derived from herpes simplex virus (HSV); empty square, β -globin TATA box; β -globin and REF-globin, reporter genes based on the rabbit β -globin gene. **b**, HeLa cells were transfected¹⁹ with 5 μ g of 4 x Sp1 (SV40) and 4 x Sp1 (HSV) reporter plasmids. In lanes 2–4, 10 μ g of the indicated RPII215-CTD construct was added to the transfection mix. In lane 1, equimolar amounts of empty expression vector were added and the total amount of plasmid DNA was adjusted to 20 μ g by adding pUC18. Lane M, HpaII digested pBR322 DNA.

METHODS. To construct 4 x Sp1 (HSV), four copies of the Sp1 binding site (TGGGCGGGGCCGGGGA) derived from HSV were cloned into the SalI site (–42) of the standard OVEC-REF construct. To construct 4 x Sp1 (SV40), a 90-bp NcoI–SphI fragment from SV40, including two copies of the 21-bp repeat (TCCCGCCCTAACTCCGCCCA), was cloned into the SalI site (–42) of OVEC-0 (ref. 22). To allow for the synthesis of the α -amanitin-resistant largest subunit, HeLa (or COS-1) cells were transfected and grown for 16 h before α -amanitin was added to the medium for an additional 35 h. The decline in background transcription in cells transfected with only the reporter plasmid indicates that the half-life of the reporter mRNA is short compared to the time of incubation in the presence of α -amanitin. Transient expression of all pol II clones in the absence of α -amanitin allows sufficient expression of the mutant form to enable its transcription capability to be assessed in the presence of α -amanitin. Trypan-blue staining reveals no increase in cell death in $\Delta 5$ transfected cultures compared to the $\Delta 31$ and WT transfected cultures after 40 h of amanitin treatment; growth curves were the same in the three transfected cultures during the time of the experiment (data not shown).

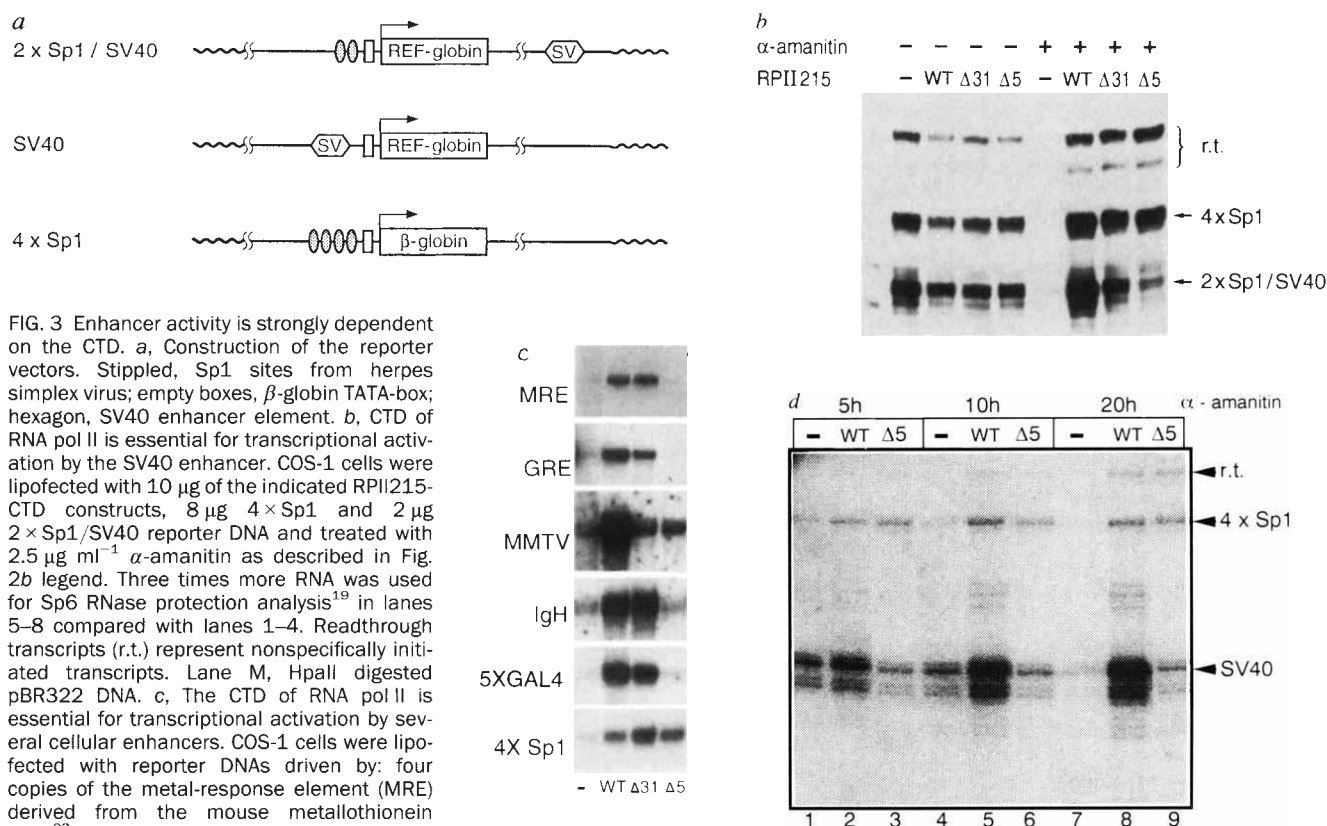


FIG. 3 Enhancer activity is strongly dependent on the CTD. **a**, Construction of the reporter vectors. Stippled, Sp1 sites from herpes simplex virus; empty boxes, β -globin TATA-box; hexagon, SV40 enhancer element. **b**, CTD of RNA pol II is essential for transcriptional activation by the SV40 enhancer. COS-1 cells were lipofected with 10 μ g of the indicated RPII215-CTD constructs, 8 μ g 4 $\times$ Sp1 and 2 μ g 2 $\times$ Sp1/SV40 reporter DNA and treated with 2.5 μ g ml⁻¹ α -amanitin as described in Fig. 2b legend. Three times more RNA was used for Sp6 RNase protection analysis¹⁹ in lanes 5–8 compared with lanes 1–4. Readthrough transcripts (r.t.) represent nonspecifically initiated transcripts. Lane M, HpaII digested pBR322 DNA. **c**, The CTD of RNA pol II is essential for transcriptional activation by several cellular enhancers. COS-1 cells were lipofected with reporter DNAs driven by: four copies of the metal-response element (MRE) derived from the mouse metallothionein gene²³, four copies of the glucocorticoid response element (GRE)²⁴, the mouse mammary tumour virus LTR (MMTV, 'MTV- β ')²⁵, the immunoglobulin heavy-chain enhancer (IgH), or five copies of the GAL4 factor binding site (GAL4)²⁶. The control 4 $\times$ Sp1 is shown below. For the GRE and GAL4 enhancers, plasmids expressing the constitutive glucocorticoid receptor variant²⁷ or GAL4-VP16²⁶ were cotransfected. **d**, Time course of α -amanitin incubation. COS-1 cells were lipofected with 1 μ g SV40, 9 μ g 4 $\times$ Sp1 reporter and 10 μ g RPII215-CTD-constructs as in **b**. Cells were collected after the time indicated (in hours) of α -amanitin incubation and 6 μ g RNA was analysed by S1 mapping²². The figure was printed directly from a phosphorimager file.

Activators like the glutamine-rich Sp1 protein interact with the transcription factor TFIID to stabilize preinitiation complexes¹⁵. An alternative activation mechanism involves a mediator, which may draw pol II into the preinitiation complex⁵. A recently identified yeast RNA polymerase II holoenzyme may contain this mediator bound to the CTD^{4–6}. In yeast, activator binding sites are typically found within 1 kilobase of the initiation site¹⁶, whereas mammalian transcription units may contain, in addition to such promoter proximal sequences, enhancer elements located many kilobases from the initiation site in either the 5' or 3' direction^{17–19}. Our results indicate that transcription

METHODS. To construct 2 $\times$ Sp1/SV40, the 196 bp *Eco*RI SV40 enhancer fragment was cloned in the 1.5-kb downstream *Eco*RI site of the 2 $\times$ Sp1-OVEC-REF construct. For SV40, the 196-bp SV40 enhancer fragment was cloned into the *Sal*I site of the OVEC-REF reporter plasmid. The IgH reporter was constructed by inserting the mouse IgH heavy-chain enhancer fragment (Asp 718/*Eco*RI) from pB1H.3.1 (ref. 28) into the *Eco*RI site of OCTA(1)²⁹. Because the SV40 enhancer stimulates transcription to a higher level than the Sp1 promoter, we reduced the amount of enhancer-driven reporter plasmid 4- or 5-fold compared to the 4 $\times$ Sp1 promoter plasmid in cell transfection such that similar levels of transcription were obtained.

activation by promoter-proximal binding proteins like Sp1 and enhancer-binding proteins may take place through different mechanisms, although we cannot exclude a quantitative difference in the number of CTD repeats required for the two processes. Yeast GAL4 activation has been demonstrated *in vitro* using an RNA polymerase II holoenzyme consisting of pol II and a high-molecular-weight mediator complex^{4–6}. Several components of the mediator were originally identified as suppressors of CTD truncation mutations, indicating that the CTD and the mediator function in the same pathway⁴. It will be of interest to see whether a similar mediator complex functions in mammalian cells to affect enhancer function. □

Received 8 July 1994; accepted 23 February 1995.

- Corden, J. L., Cadena, D. L., Ahearn, J. Jr & Dahmus, M. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7934–7938 (1985).
- Allison, L. A., Moyle, M., Shales, M. & Ingles, C. J. *Cell* **42**, 599–610 (1985).
- Corden, J. L. & Ingles, C. J., in *Transcriptional Regulation* (eds McKnight, S. L. & Yamamoto, K. R.) 81–108 (Cold Spring Harbor Laboratory Press, Plainview, New York, 1992).
- Thompson, C. M., Koleske, A. J., Chao, D. M. & Young, R. A. *Cell* **73**, 1361–1375 (1993).
- Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H. & Kornberg, R. D. *Cell* **77**, 599–608 (1994).
- Koleske, A. J. & Young, R. A. *Nature* **368**, 466–469 (1994).
- Allison, L. A. & Ingles, C. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2794–2798 (1989).
- Scafe, C. et al. *Nature* **347**, 491–494 (1990).
- Bartolomei, M. S., Corden, J. L. *Molec. cell Biol.* **7**, 586–594 (1987).
- Müller-Sturm, H. P., Sogo, J. M. & Schaffner, W. *Cell* **58**, 767–777 (1989).
- Zehring, W. A. & Greenleaf, A. L. *J. biol. Chem.* **265**, 8351–8353 (1990).
- Buratoski, S. & Sharp, P. A. *Molec. cell Biol.* **10**, 5562–5564 (1990).
- Bartolomei, M. S., Halden, N. F., Cullen, C. R. & Corden, J. L. *Molec. cell Biol.* **8**, 330–339 (1988).
- Gerber, H. P., Georgiev, O., Harshman, K. & Schaffner, W. *Nucleic Acids Res.* **20**, 5855–5856 (1992).
- Gill, G., Pascal, E., Tseng, Z. H. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **91**, 192–196 (1994).

- Guarente, L. *Cell* **36**, 799–800 (1984).
- Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299–308 (1981).
- Moreau, P. et al. *Nucleic Acids Res.* **9**, 6047–6068 (1981).
- Seipel, K., Georgiev, O. & Schaffner, W. *EMBO J.* **11**, 4961–4968 (1992).
- Rusconi, S., Severne, Y., Georgiev, O., Galli, I. & Wieland, S. *Gene* **89**, 211–221 (1990).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
- Westin, G., Gerster, T., Mueller, M. M., Schaffner, G. & Schaffner, W. *Nucleic Acids Res.* **15**, 6787–6798 (1987).
- Radtke, F. et al. *EMBO J.* **12**, 1355–1362 (1993).
- Schatt, M. D., Rusconi, S. & Schaffner, W. *EMBO J.* **9**, 481–487 (1990).
- Wieland, S., Gail, I., Schatt, M., Severne, Y. & Rusconi, S. in *Activation of Hormone and Growth Factor Receptors* (eds Sekeris, C. E. & Alexis, M.) 215–225 (Kluwer Academic, Dordrecht, 1990).
- Seipel, K., Georgiev, O. & Schaffner, W. *EMBO J.* **11**, 4961–4968 (1992).
- Severne, Y., Wieland, S., Schaffner, W. & Rusconi, S. *EMBO J.* **7**, 2503–2508 (1988).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729–740 (1983).
- Muller, M. M., Ruppert, S., Schaffner, W. & Mathias, P. *Nature* **336**, 544–551 (1988).

ACKNOWLEDGEMENTS. We thank M. Birnstiel for the pVA1 plasmid, C. Hovens and P. J. Mitchell for critical reading of the manuscript, M. Gstaiger for discussion, S. Rusconi for GR plasmids, S. Wieland for the CMV-OVEC construct, H.-P. Müller for initiating the project, and F. Ochsenbein for artwork.